UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA



DIVISIÓN DE INGENIERÍA MECÁNICA E INDUSTRIAL

ÁLGEBRA LINEAL
Y SUS APLICACIONES

IDALIA FLORES DE LA MOTA





UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA



ÁLGEBRA LINEAL
Y SUS APLICACIONES

IDALIA FLORES DE LA MOTA

DIVISIÓN DE INGENIERÍA MECÁNICA E INDUSTRIAL

FLORES DE LA MOTA, Idalia.
Álgebra lineal y sus aplicaciones
México, Universidad Nacional Autónoma de México,
Facultad de Ingeniería, 2021, 345 p.

ISBN 978-607-30-4139-3

ÁLGEBRA LINEAL Y SUS APLICACIONES

Marzo, 2021

FACULTAD DE INGENIERÍA
http://www.ingenieria.unam.mx/

ISBN 978-607-30-4139-3



Unidad de Apoyo Editorial
Cuidado de la edición: María Cuairán Ruidíaz
Diseño de forros: Nismet Díaz Ferro

Contenido

PRÓLOGO

Álgebra *lineal y sus aplicaciones* tiene como antecedente inmediato los apuntes de Álgebra *lineal* publicados en 1991 cuando la autora impartía dicha materia en el Posgrado de Ingeniería que entonces era DEPFI. Esta nueva publicación conserva la estructura y la teoría de la anterior, incluyendo aplicaciones tomadas de otros autores, así como ejercicios propuestos al final de cada capítulo. En las notas históricas se agregaron ilustraciones y algunos detalles más que las hacen más interesantes para el lector. El orden de los temas es común a los libros de álgebra lineal, pero considerando que se aborda cada capítulo conforme al grado de dificultad de los temas tratados.

El Álgebra Lineal es una estructura matemática de suma importancia en nuestro medio; ya que no solo trata de resolver problemas de Investigación de Operaciones, también es una herramienta importante en todas las áreas de la ingeniería. Muchas veces los textos que tratan sobre esta disciplina son muy teóricos y se limitan a resolver unos cuantos ejemplos abundando en la demostración de teoremas, lo cual no es nuestro propósito. Más bien, se tiene como objetivo resolver una gran cantidad de ejemplos sobre todo de ingeniería y efectuar las demostraciones de los teoremas más relevantes. En cada capítulo se presentan aplicaciones de los temas abordados ahí con la finalidad de que el lector tenga conocimiento del potencial del tema estudiado. Bajo esta perspectiva, se presentan estos apuntes donde lo que se busca es tener un apoyo didáctico para la clase y como un complemento de la bibliografía sugerida para el curso, en este sentido es importante mencionar que no pretenden sustituir a los libros de álgebra lineal, son simplemente un apoyo para la materia. No se presenta ningún complemento computacional ya que este va cambiando y desarrollándose muy rápidamente.

Los presentes apuntes van dirigidos a estudiantes de licenciatura y posgrado que requieren de un cierto grado de conceptualización y una buena dosis de ejemplos y aplicaciones. También al final de cada capítulo se presentan notas

históricas que buscan dar al lector una idea más completa de las matemáticas, viéndolas desde una perspectiva histórica, inmersas en la realidad de la época. Parafraseando a Isaac Asimov: "La ciencia gana realidad cuando es visualizada no como una abstracción, sino como la suma concreta de los científicos, pasados y presentes, vivos y muertos. No hay estamentos en ciencia, ni una observación, ni un pensamiento que exista por sí mismo. Cada uno de ellos es producto de un duro esfuerzo de algún hombre, y a menos que conozcáis al hombre y el mundo con el cual trabajaba, los supuestos que él aceptaba como verdades, los conceptos que él consideraba insostenibles, no podréis comprender sus afirmaciones, u observaciones, o pensamiento".

Otro objetivo al publicar este trabajo fue la celebración del 50 aniversario de la Maestría en Investigación de Operaciones, que no ha visto la luz hasta ahora debido a problemas técnicos que se presentaron para su revisión en el tiempo previsto.

Quiero agradecer el apoyo en revisión y correcciones a la Maestra María Cuairán Ruidíaz, jefa de la Unidad de Apoyo Editorial de la Facultad de Ingeniería de la UNAM, por el respaldo brindado y por la colaboración para la realización de la edición de esta obra.

A la DGAPA (Dirección General de Asuntos del Personal Académico) de la UNAM por el apoyo económico proporcionado para la elaboración de este material, a través del proyecto PAPIME PE-107116.

Agradezco a Olivia Sashiko Shirai Reyna, José Antonio Sánchez Calderón y Rosalío Arteaga Montiel que hicieron revisiones y aportaron conocimientos en el desarrollo de estos.

Capítulo 1
MODELOS MATEMÁTICOS

"La formulación de un problema es a menudo
más importante que su solución"
A. Einstein

1.1. INTRODUCCIÓN

Un modelo es una representación de un problema o situación de la realidad. Esta representación la hacemos mediante diversos objetos o símbolos a través de un proceso de abstracción, que consiste en tomar de la realidad los elementos más importantes que intervienen en el problema y desechar todos aquellos que consideramos que no juegan un papel determinante en el mismo, estableciendo con precisión cuáles son las distintas relaciones que guardan entre sí dichos elementos. Una vez establecidas estas relaciones, podemos manipular los elementos del modelo en la búsqueda de una posible solución; o demostrar que tal solución no existe.

Se considera que las funciones de un modelo son la predicción y la comparación para proporcionar una forma lógica de predecir los resultados que siguen las acciones alternativas, e indicar una preferencia entre ellas. Aunque este uso de los modelos es importante, no es, de ninguna manera, su único propósito; la construcción de modelos proporciona una manera sistemática, explícita y eficiente para que un grupo de expertos y aquellos que toman las decisiones centren su juicio e intuición.

Uno de los principales elementos para resolver un problema es la construcción y el uso de un modelo. Dicho modelo puede tomar muchas formas, pero una de las más útiles, y ciertamente la que más se usa, es la matemática; lamentablemente, no siempre es posible crear un modelo matemático en un sentido riguroso y estricto. Al estudiar la mayoría de los sistemas militares e industriales, podemos definir objetivos, especificar restricciones y discernir si el diseño

sigue las leyes de ingeniería y/o economía. Las relaciones esenciales se pueden descubrir y representar matemáticamente, de una manera u otra.

Representar algún objeto, sistema o idea con un modelo es tan general que es difícil clasificar todas las funciones que cumplen los modelos, algunas de ellas son:

a) Una ayuda para el pensamiento
b) Una ayuda para la comunicación
c) Para entrenamiento e instrucción
d) Una herramienta de predicción
e) Una ayuda para la experimentación

En ingeniería, los modelos sirven como ayuda para diseñar nuevos sistemas o mejorar los existentes, mientras que en ciencias sociales y en economía, explican los sistemas existentes.

1.2. CARACTERÍSTICAS DE LOS MODELOS

Empezaremos nuestro análisis con un problema hipotético:

> **El viejo y el lobo**
> “A orillas del río Balsas, un viejo llevaba un lobo, una oveja y una paca de alfalfa. El viejo quiere cruzar el río con todas sus pertenencias, pero lo único que tiene para cruzarlo es una barca donde solo cabe él y una de sus pertenencias de tal manera que tiene que llevarlas a la otra orilla una por una. Pero si deja al lobo con la oveja, este se la comería y si deja a la oveja con la alfalfa, esta se la comería”. ¿Cómo puede el viejo llegar a la otra orilla del río con todas sus pertenencias?

Antes de resolver el problema, debe estar claro que podemos considerarlo como un sistema. Empecemos por diferenciar las relaciones que guardan entre sí los elementos del sistema:

1. En la barca solo cabe el viejo con una de sus pertenencias (lobo, oveja, alfalfa)
2. El lobo no se puede quedar solo con la oveja
3. La oveja no se puede quedar sola con la alfalfa

Una vez enumeradas las principales relaciones de interdependencia del sistema, podemos intentar diseñar una posible solución.
A continuación, se enuncian los pasos a seguir para una solución satisfactoria del problema:

1. El viejo cruza el río con la oveja
2. El viejo regresa solo
3. El viejo cruza el río con el lobo
4. El viejo regresa con la oveja
5. El viejo cruza el río con la alfalfa
6. El viejo regresa solo
7. El viejo cruza con la oveja

Ahora pasemos a verificar si esta solución cumple las relaciones de interdependencia del sistema; esto es, hay que verificar que en ningún momento se pueda comer el lobo a la oveja ni la oveja a la alfalfa y que al final estén todos al otro lado del río. Una vez realizado lo anterior podemos, ahora sí, estar seguros de que la solución propuesta es adecuada.

Hasta ahora, sin decirlo, hemos seguido un método para resolver el problema; sin embargo, es importante que conozcamos otros más y sepamos diferenciarlos. El problema anterior lo pudimos resolver de manera sencilla en forma mental, para lo cual lo único que tuvimos que hacer fue imaginar el sistema con sus relaciones de interdependencia y a través de estas representaciones mentales, encontramos la solución; en otras palabras, elaboramos un modelo "mental" del sistema. A este método de solución le llamaremos RESOLUCIÓN MENTAL. Sin embargo, no todos los problemas se pueden resolver tan fácilmente, existen problemas mucho más complicados para cuya solución necesitamos de otros métodos.

Regresemos al problema inicial y con la finalidad de dejar más claro el razonamiento que hemos seguido, resolvamos las siguientes preguntas:

1. ¿Cuántas veces tuvo que cruzar el viejo el río?
2. Si la oveja no se comiera la alfalfa, ¿cuántas veces cruzaría el río?
3. ¿Hay otras soluciones al problema?
   ¿Existe una solución más corta?
   Si ponemos alfalfa donde diga lobo y lobo donde diga alfalfa en nuestra solución, ¿obtenemos otra solución?

Si intercambiamos alfalfa y oveja, ¿es solución?

4. ¿Únicamente podemos resolver mentalmente nuestro problema?
5. ¿De qué otra manera, diferente a la mental, podemos resolver el problema?

Preguntas como estas, corresponden a lo que en investigación de operaciones se conoce como análisis de sensibilidad, las últimas nos sugieren que existen más métodos de solución y en caso de existir otras soluciones, la siguiente pregunta obligada es ¿cuál es la mejor solución o la solución óptima?

Otro método de solución es trasladarnos al lugar de los hechos y ahí encontrar la solución del problema. Esto es: trasladarnos al río Balsas, conocer al viejo y demás entidades del sistema, decirle al viejo cómo pasar el río, cuidando de que no haya ningún error y una vez que se encontrase en la otra orilla sano y salvo con todas sus pertenencias, regresarnos. Este método de resolver el problema lo llamaremos SOLUCIÓN DIRECTA. Obviamente, este último método presentaría dificultades y en el supuesto de nuestro problema, nuestra simple presencia modificaría el problema original del viejo.

Existe otro método que es muy utilizado en la solución de problemas y al cual le daremos el nombre de SIMULACIÓN. Como su nombre lo indica, este método consiste en hacer una representación del sistema en cuestión, con personas, objetos, etc., e interactuar con el modelo hasta encontrar la solución, que suponemos válida para el problema real.

Finalmente, un tipo de método muy importante y muy común es aquel que utiliza símbolos escritos en un papel. A este se le ha dado el nombre de método SIMBÓLICO. Así, el método simbólico resuelve problemas a través de dibujos y diagramas (símbolos), mediante los cuales se representan las entidades del sistema, cuidando de utilizar símbolos diferentes para cada entidad distinta.

Por ejemplo, en nuestro problema del viejo podríamos representar los elementos del problema de la siguiente manera:

	Representa al viejo		Representa la alfalfa
	Representa al lobo		Representa la barca
	Representa la oveja		Representa al río

FIGURA 1.1 Representación del problema

De esta manera podríamos representar cada uno de los pasos a seguir en la resolución del problema; sin embargo, esto entraña el problema de que algunos de nosotros no sabemos dibujar bien, por lo que el modelado sería muy complicado. Así que busquemos símbolos más sencillos de dibujar, por ejemplo, sus iniciales (esto es: V, L, O, A, B, para designar al viejo, al lobo, la oveja, la alfalfa y la barca) y usemos dos líneas verticales para representar al río y una flecha para indicar la dirección en que se efectúa el cruce.

La solución de nuestro problema quedaría:

Todos los personajes en una orilla del río	VLOA	B	
1. El viejo cruza el río con la oveja	LA LA	BVO → B	VO
2. El viejo regresa solo	LA LAV	← VB B	O O
3. El viejo cruza el río con el lobo	A A	BVL → B	O VLO
4. El viejo regresa con la oveja	A AOV	← BVO B	L L
5. El viejo cruza el río con la alfalfa	O O	BVA → B	L VAL
6. El viejo regresa solo	O	← BV	AL
7. El viejo cruza con la oveja		BVO → B	AL VOAL

De esta forma, el viejo y sus pertenencias quedan a la otra orilla del río. Si observamos detenidamente nuestro modelo simbólico, notamos que se puede simplificar más. Se deja de ejercicio esa simplificación.

En conclusión, se puede decir que en cada uno de los modelos se sustituye cada elemento de la situación real por otro que lo represente. Por otro lado, las diferencias entre los distintos modelos es que dado el nivel de abstracción cada vez se parecen menos al problema real, pero se gana en generalidad pues se puede aplicar el modelo a otros problemas para su solución.

1.3. LA ABSTRACCIÓN

En todos los métodos de solución de problemas mencionados, se tiene en común la utilización de un modelo. En el método mental, nuestro modelo fueron imágenes en nuestro cerebro. En el método directo nuestro modelo fue la realidad misma (sin lugar a duda el mejor modelo). En el de simulación, nuestro modelo consistió en representaciones del problema por medio de personas o cosas, suponiendo que cada una de ellas guardaba entre sí las mismas relaciones que las del problema original y finalmente, en el método simbólico nuestro modelo consistió en símbolos en el papel.

Cabe resaltar que en los últimos dos métodos (de simulación y simbólico) utilizamos distintos modelos.

Fijaremos nuestra atención en los modelos simbólicos: primero, representamos la situación del problema y sus personajes por medio de dibujos, por ello nuestro modelo es de fácil comprensión para todos, pero tiene la desventaja de ser muy laborioso, resulta casi imposible de construir a medida que el número de variables y entidades aumenta.

Como segunda opción, utilizamos letras para representar a nuestros personajes (letras relacionadas con ellos). Este modelo tiene la ventaja de que es fácil de elaborar y la desventaja de que, para entenderlo, necesitamos más información sobre el problema y requerimos un mayor esfuerzo mental, esto debido a que nuestro modelo no se parece a la realidad.

1.4. MODELOS MATEMÁTICOS

Ya que se está hablando de abstracción, una manera de enfrentar y resolver problemas como el planteado y algunos otros relativos al área de las ingenierías y las ciencias exactas son los modelos matemáticos.

En los capítulos de este libro nos enfocaremos a modelos que se refieren a relaciones y ecuaciones lineales, así como a algunas aplicaciones que publicadas en libros de álgebra lineal.

El álgebra lineal es una rama de las matemáticas que considera una serie de fenómenos problemas cuyo comportamiento es lineal, es decir se puede expresar como una recta, un plano, y conjunto de ellos y sus extensiones. Es un área que se considera sumamente abstracta y por lo mismo no es muy popular entre la gente ni en las universidades.

Por todo lo anterior se tiene como un objetivo general de este material, mostrar los conceptos y la teoría, así como ejemplos y ejercicios para comprender sus mecanismos y vincularlos en la medida de lo posible a aplicaciones reales y viceversa, como es que a través de problemas concretos se encuentra un uso de estos modelos.

Los modelos matemáticos se abordarán en los subsecuentes capítulos, en el presente se da una visión general de modelado. Como veíamos con nuestro caso hipotético otra manera de resolver problemas es usando modelos de simulación, cuya diferencia con los modelos de álgebra lineal radica en su mayor flexibilidad, así como aleatoriedad.

Un enfoque diferente a la representación por medio de modelos de sistemas (complejos) consiste en utilizar la simulación. Los modelos de simulación difieren de otros modelos en que las relaciones entre la entrada y la salida no se indican en forma explícita.

Un modelo de simulación divide al sistema representado en módulos básicos conectados por relaciones lógicas bien definidas (en la forma si/entonces). Por tanto, partiendo del módulo de entrada, las operaciones de cálculo pasarán de un módulo a otro hasta que se obtenga un resultado de salida. Los modelos de simulación, en comparación con otros modelos, como ya se decía anteriormente ofrecen una mayor flexibilidad en la representación de sistemas complejos.

1.5. MODELOS MATEMÁTICOS Y SIMULACIÓN

Pero aquí cabe la pregunta: ¿Qué es la simulación?

Un concepto intuitivo es: simular, es representar, fingir, actuar. En la ciencia, la industria y la educación no es algo distinto: la simulación es una técnica de investigación o enseñanza, que reproduce en forma semejante o aproximada los eventos reales y los procesa con ciertas condiciones de prueba, definidas con anterioridad. Desarrollar simulaciones de este tipo requiere de procesos matemáticos, que en algunos casos son complejos. Inicialmente, debe especificarse un conjunto de reglas, relaciones y procedimientos operativos. La interacción de estos fenómenos crea nuevas situaciones o nuevas reglas que evolucionan al desarrollarse la simulación.

La forma de implementar la simulación va desde objetos muy sencillos como papel y lápiz, hasta sofisticadas representaciones en computadora, con sistemas interactivos de entornos casi reales.

Una definición actual y adecuada de simulación puede ser la siguiente:

> ***Definición 1.1***
> La simulación es una técnica numérica para conducir experimentos en una computadora digital, haciendo uso de gráficos, animación y otros dispositivos tecnológicos; lo cual involucra ciertos tipos de modelos matemáticos y lógicos, que describen el comportamiento de un sistema (o algún componente de éste) durante un cierto tiempo.

La otra pregunta es: ¿Cuándo usar simulación?

La simulación es una de las técnicas administrativas más frecuentemente usadas, y todo parece indicar que su popularidad va en aumento. Para analizar las razones de su uso, es interesante explorar las alternativas existentes a la simulación, es decir, los diferentes métodos que pueden usarse para resolver el mismo problema:

1. Uso de algún otro tipo de modelo matemático de tipo analítico;
2. Experimentación directa con el modelo real o con un prototipo de éste;
3. Uso de la experimentación y la intuición.

La figura 1.2 muestra un diagrama de decisión para el uso de modelos.

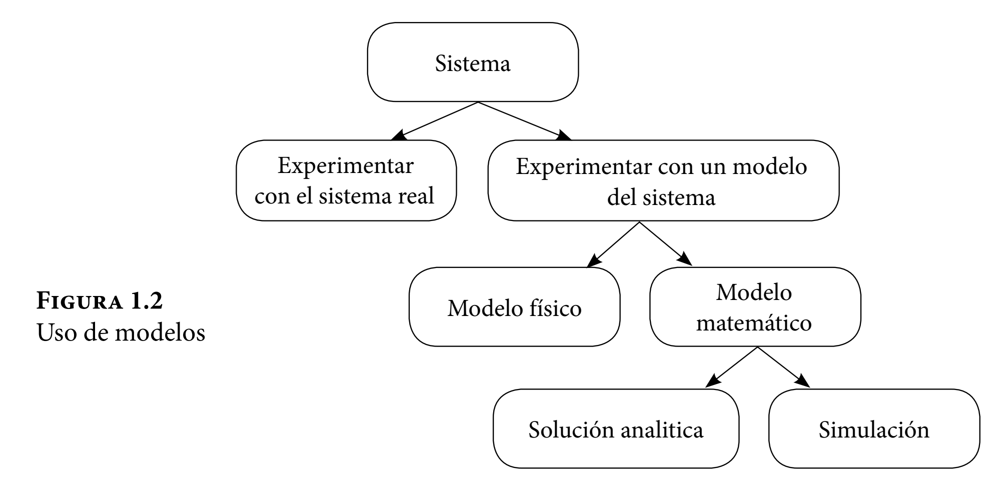


Figura 1.2
Uso de modelos

En realidad, si la solución analítica es relativamente sencilla, siempre será preferible a la simulación, ya que se considera al modelo general. Sin embargo, el problema es que existen muchos sistemas que no generan problemas sencillos de resolver, en este caso se recurre a la simulación. Por ejemplo, se tienen las colas o líneas de espera que involucran procesos aleatorios distribuidos en una serie de componentes del sistema, los modelos de inventarios, de recursos compartidos, de pronósticos de series de tiempo, de comportamientos económicos, esquemas de producción, movimiento de vehículos y dinámicas de cruceros viales entre otros.

Otra ventaja de la simulación es que se puede experimentar sin exponer a la organización a los perjuicios de errores en el mundo real. Por ejemplo, algunos bancos han cambiado su sistema de filas múltiples a fila única sin necesidad de experimentar con los clientes, ya que esto podría tener consecuencias desagradables si no funciona como se espera.

Por otro lado, es más sencillo controlar condiciones experimentales en un modelo de simulación que en un sistema real. Podemos pensar en un modelo de un crucero vial, donde puedan analizarse diferentes sincronizaciones de semáforos sin afectar a los elementos reales, lo cual tendría un costo excesivo, que podría llegar hasta lo invaluable de una vida humana.

En un modelo de simulación es posible comprimir largos periodos de tiempo y analizar el comportamiento en forma inmediata. Podemos visualizar cómo será la población dentro de 30 años y si los servicios de transporte serán suficientes para satisfacerla.

Por supuesto, existen casos en los que el sistema que se quiere analizar ni siquiera existe, de modo que, definitivamente lo ideal será usar la simulación o algún método de tipo cualitativo. La simulación no reemplaza a otras formas de experimentación ni al juicio subjetivo, pero es una solución alternativa conveniente.

La experiencia y la intuición, así como el profundo conocimiento de los fenómenos, deberán ser ingredientes constantes para el éxito de los modelos de simulación.

Todo lo que se refiere a modelar ya sea con simulación o modelos analíticos tiene mucho que ver con los problemas en cuestión a resolver, en nuestro caso nos centraremos a problemas cuyas alternativas de solución son analíticas y no requieren el uso de simulación. Sin embargo, siempre es importante contar con más de una herramienta a la hora de proponer modelos ya que un paso importante en la solución es la validación del modelo.

1.6. RESUMEN

1. Los modelos simbólicos que utilizamos serán cada vez menos reales y más abstractos. En su elaboración efectuamos un proceso de abstracción.
2. Observe que, en los modelos que hemos construido, una vez que elegimos un objeto o símbolo para representar una entidad, éste permanece sin cambiar a lo largo de todo el proceso. Esta es una característica que todo modelo debe satisfacer.
3. De los métodos de solución que hemos visto, existe una gran diferencia entre el método directo y los demás. En el método directo nosotros no construimos el modelo, el modelo en este caso es la realidad misma. Además de las entidades que hemos considerado intervienen muchas más como, por ejemplo: hay rápidos o las aguas son tranquilas, el río es ancho o angosto, es de día o de noche, etc., factores que sin lugar a duda intervienen y pueden facilitar o dificultar la solución del problema. En los demás métodos nosotros hemos construido el modelo, tomando de la realidad solo las entidades que consideramos importantes y desechando las demás.
4. Otra consideración importante es que, si en nuestro modelo nos equivocamos en la búsqueda de la solución, lo único que se habrá perdido es tiempo y un pedazo de papel, mientras que un error, en el método directo, implica la pérdida de una de las pertenencias del viejo. Un error puede corregirse sin mucha dificultad en los modelos, mientras que en la realidad difícilmente podremos hacerlo, ya que la experiencia nos ha enseñado que toda acción conlleva un riesgo.
5. Se preguntarán ahora, ¿De qué nos sirve resolver un problema en un modelo si, como vimos en nuestro ejemplo, hemos olvidado algunos factores que podrían imposibilitar la solución (tal como la posibilidad de que caiga una tormenta)? A lo que podríamos contestar: la solución encontrada en nuestro modelo nos sirve de guía para la acción, para intentar solucionar felizmente un problema. Además, en nuestro modelo proponemos la solución al problema de cualquier viejo o joven, que se encuentre en esas condiciones y no solo en el río Balsas, sino en cualquier otro rio; proponemos la solución no de uno, sino de muchos problemas. Por esto la solución en un modelo es más general.
6. Podemos afirmar que entre más abstracto es un modelo, más generales son sus resultados; es decir, se aplican a un mayor número de casos, pero también hay que recordar que para elaborar e interpretar un modelo, entre más abstracto es más difícil hacerlo.

7. Finalmente, no hay que perder de vista que existen diferencias sustanciales entre resolver un problema en un modelo y resolverlo en la realidad. Lo que significa que no debemos ser tan mecánicos en la aplicación a la realidad de una solución obtenida a través de un modelo, ya que no podemos olvidar que en la realidad la situación es cambiante, por consiguiente, elementos que en un principio desechamos al elaborar el modelo pudieron haberse convertido en determinantes, modificando con ello las condiciones reales del problema, lo cual imposibilita la aplicación de la solución encontrada.

1.7. NOTAS HISTÓRICAS

(Tomado de *La historia de las matemáticas: de dónde vienen y hacia dónde se dirigen*, Benjamín Galán Atienza, 2012)

Las matemáticas son tan antiguas como el propio conocimiento humano. Se puede apreciar en los diseños prehistóricos de utensilios de cerámica, pinturas en los que se aprecia la utilización de geometría.

También sabemos que el método de cálculo de los primitivos consistía en el uso de los dedos de las manos para contar y eso se ve reflejado en los tipos de sistemas numéricos cuyas bases son de cinco y diez.

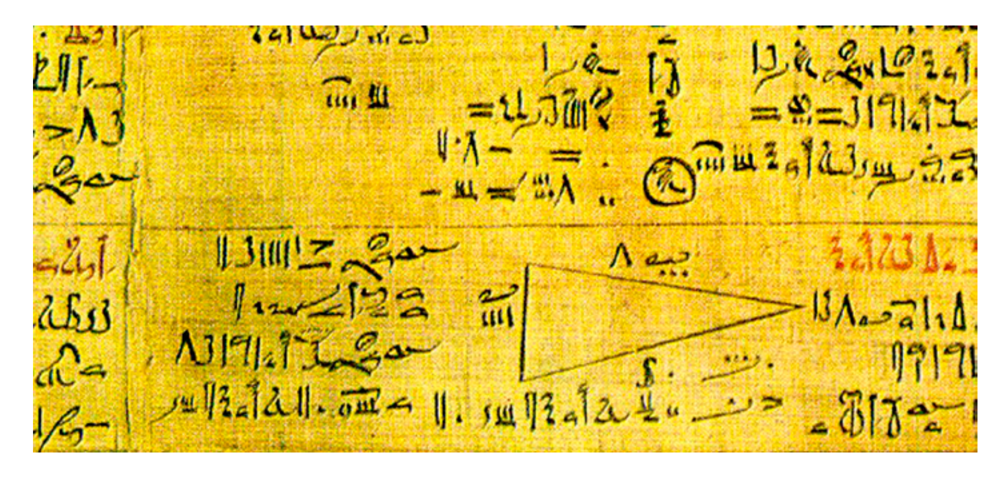
Más tarde empezaron las civilizaciones a tener un pensamiento más profundo sobre las matemáticas.

Las primeras civilizaciones de la que se tiene constancia de la utilización de las matemáticas para su desarrollo fueron las civilizaciones egipcia y babilónica.

Egipto y Babilonia. Los primeros conocimientos de referencias de utilización de matemáticas en una cultura datan del 3.000 antes de Cristo. Empezaron a surgir en la zona de Egipto y Babilonia y posteriormente se fueron expandiendo por todo el mundo. Esta cultura utilizaba las matemáticas como una pura aritmética. Se preocupaban un poco de la forma de los objetos y los diferentes tipos de geometría, pero no utilizaban demostraciones matemáticas y tampoco tenían el concepto de la creación de postulados, como referencia para avanzar en la ciencia. Son unas matemáticas prácticas para los problemas de su sociedad.

Por otro lado, en oriente las matemáticas también estaban teniendo un papel importante en el desarrollo de las civilizaciones. Gracias a las rutas comerciales, se conocían los métodos matemáticos en muchas partes del mundo. China,

como relataremos a continuación también fue una civilización basada en el comercio y desarrolló las matemáticas para así poder potenciar, entre otras muchas cosas, su crecimiento comercial.



Papiro de Rhind - Autor Ahmes – Fecha: aprox. 1650 a.C.

China. El inicio de las matemáticas en el pueblo chino se puede comparar en antigüedad a las civilizaciones de Egipto y Mesopotamia. Uno de los primeros descubrimientos que se conoce del pueblo chino, es el descubrimiento de las horas solares. Este hecho viene incluido en la obra matemática llamada Chou Peique data del 1200 a. C. Es la mayor obra matemática china y está formada por nueve libros o capítulos.
Está compuesta por pergaminos y escritos independientes y recogen todos los temas importantes para su pueblo planteados en 246 problemas específicos. Este planteamiento de la resolución de los problemas, también lo realizaron el pueblo egipcio y el pueblo babilónico. El Chou Pei contenía problemas sobre agricultura, ingeniería, comercio y también aparece en el capítulo 8 un logro importante de cómo resolver ecuaciones lineales, y sistemas complejos de cuatro ecuaciones con cinco incógnitas y ecuaciones indeterminadas.

Los chinos al igual que el resto de las culturas, necesitaban resolver los problemas de la vida diaria y sus matemáticas reflejaban el modo de vida que tenían. Sus actividades principales eran la agricultura, la ingeniería poco avanzada, y adaptaron las matemáticas para resolver problemas de impuestos. También utilizaron las matemáticas para problemas de ecuaciones, así pudiendo resolver teoremas como las propiedades de los triángulos rectángulos.

Utilizaban un sistema de numeración con operaciones semejantes a otras culturas. También conocían los números negativos, pero no los aplicaban a las soluciones de las ecuaciones y no los reconocían como resultados viables.

Uno de los descubrimientos matemáticos más importantes del pueblo chino fue el método para resolver ecuaciones lineales. Inventaron el "tablero de cálculo" que descompone por colores los números positivos y los números negativos y se utilizaba de una forma similar al ábaco.

En la siguiente imagen, se puede apreciar en la esquina inferior izquierda, la distribución por colores de los números positivos y negativos, y como lo representaban en sus manuscritos.

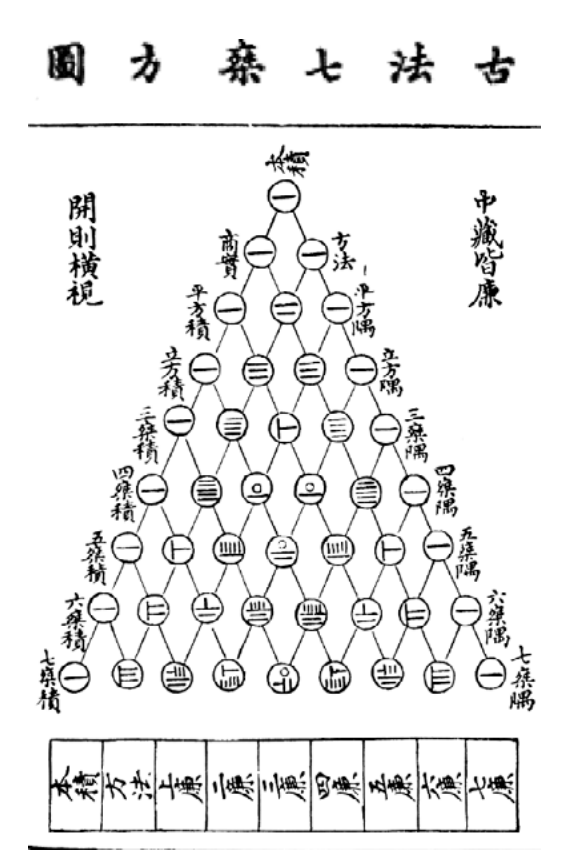


Tablero de Cálculo.
Obra Chou Pei autor
desconocido, S III a.

Chou Shi Hié desarrolló el método algebraico en la edad media, que permitía encontrar raíces enteras y racionales, y aproximar decimalmente ecuaciones de este tipo:

$$Pn(x) = a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + a_0$$

Otro gran logro fue el triángulo de Yang hui publicada en la obra Si Yuan Yu Jian 1320, que consistía en la suma de progresiones y la combinatoria, y se construyó el denominado "espejo precioso" que hoy se menciona como triángulo de Pascal.



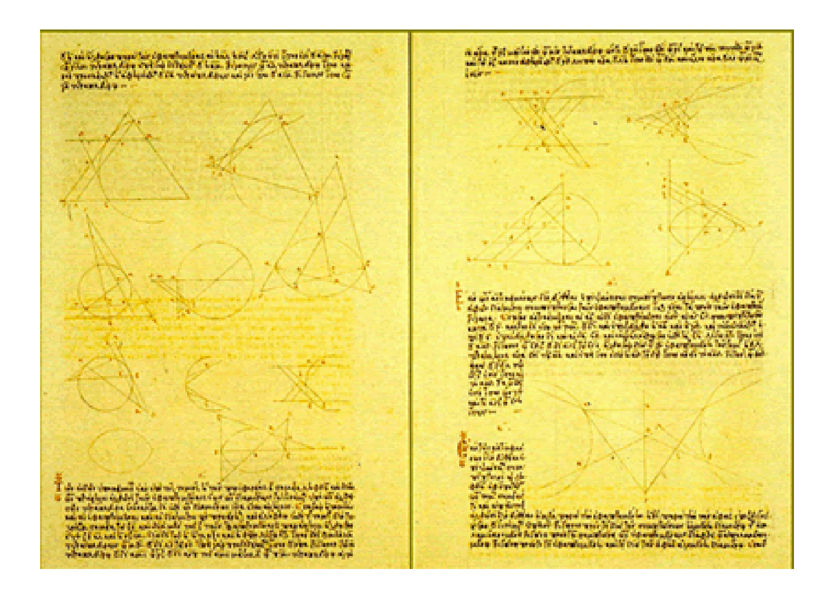
Triángulo de Yáng
Huī, 1238–1298

Se puede apreciar como el pueblo chino, un pueblo que basa su sociedad en el comercio adapta las matemáticas a ello y no tanto a otros aspectos que engloba esta materia. Durante estos procesos de evolución de la matemática china, la matemática babilónica y egipcia se iban expandiendo en los territorios próximos, influyendo en matemáticas importantes como la griega, gracias a las rutas comerciales en las cuales la evolución de los pueblos sufre puntos de unión.

Grecia. Los griegos dieron un paso que revolucionó el concepto de matemáticas y se adaptó al mundo actual. Fue la primera civilización en la que se estructuran las matemáticas a partir de definiciones, axiomas y demostraciones.

Grecia tuvo tres principales investigadores dedicados a la geometría. Fueron Euclides, Arquímedes y Apolonio y consiguieron revolucionar la geometría tal y como hoy la conocemos. También se dedicaron a la astronomía e iniciaron estudios muy reconocidos. En el ámbito de la astronomía, los griegos utilizaron el sistema babilónico de fracciones y realizaron las tablas de cuerda de un

círculo. Estas tablas relacionan el radio del círculo con la longitud de este en función del ángulo y observaron la relación que se producía. Estas tablas fueron las bases hacia el futuro de la trigonometría. Las matemáticas griegas fueron bastante más sofisticadas que las desarrolladas por otras culturas, debido a ello y a su proximidad con el resto de Europa influyeron en todo el mundo. Más tarde serían un modelo por seguir en la Edad Media, siguiendo un razonamiento inductivo establecido por reglas, definiciones y teoremas.



Archimedis opera; Apollonii Pergaei conicorum libri IIII

India. En los siglos I al VIII es cuando más se desarrollaron las matemáticas hindúes. Al ser una cultura muy religiosa, utilizaban las matemáticas frecuentemente con el fin de conseguir crear unos monumentos arquitectónicos de gran importancia y la realización de templos para adorar a sus dioses. Se vio reflejada la utilización de sistemas decimales de numeración como el resto de culturas anexas a la suya. Hay cuatro matemáticos indios que destacaron sobre el resto en aquel momento. Son Aryabhata (476-550 d.C.), Brahmagupta (598–660 d.C.), Mahavira (s. IX) y Bhaskara Akaria (s.XII). Lo más relevante de esta cultura es que utilizan reglas aritméticas para el cálculo, empiezan a utilizar los números negativos y el cero, y aceptan los números irracionales como soluciones correctas.

También consiguieron resolver las ecuaciones tanto lineales como cuadráticas y las raíces las consideraban como deudas. Los indios crearon los métodos

para resolver las ecuaciones llamadas diofánticas. Por otro lado, los árabes también en esta época tenían una evolución matemática considerable.

Arabia. Se suele creer que los números conocidos como números árabes son de esa zona geográfica, pero en realidad son hindúes. Los árabes en esta época estaban en plena expansión conquistando el mundo con la religión musulmana. Llegando así a la península Ibérica y hasta los límites de China. Debido a estas colonizaciones, el pueblo árabe iba adquiriendo la ciencia de los diversos pueblos a los que conquistaba y la hacía suya como tal.

El sistema numérico de los hindúes era de un tipo posicional y cada número tiene diferente valor en función de la posición que ocupe. Los árabes evolucionaron el sistema de los hindúes sobre las posiciones decimales y lo adaptaron a las fracciones. En la imagen siguiente se puede observar cómo evoluciona la tipografía numérica a lo largo de la historia.

Babilonia	𒐕	𒐖	𒐗	𒐘	𒐙	𒐚	𒐛	𒐜	𒐝	𒌋	
Egipto	I	II	III	IIII	IIIII	IIIIII	IIIIIII	IIIIIIII	IIIIIIIII	∩	
Grecia	A	B	Γ	Δ	E	F	Z	H	Θ	I	
Roma	I	II	III	IV	V	VI	VII	VIII	IX	X	
China Antigua	一	二	三	四	五	六	七	八	九	十	
Maya	•	••	•••	••••	—	• —	•• —	••• —	•••• —	=	
India	१	२	३	४	५	६	७	८	९	१०	०
Arabicos siglo15	1	2	3	ꝸ	4	6	Λ	8	9	10	0
Actuales	1	2	3	4	5	6	7	8	9	10	0

Numeraciones históricas

Los árabes introducen por fin los números tal y como los conocemos ahora según la posición. El 5 vale cinco en unidades, en decenas un 5 significaría 50, en centenas 500 y así sucesivamente pero este tipo de numeración no la utilizaban pueblos como el griego o el babilónico, sino que fue adoptada por los árabes que se dieron cuenta que el sistema hindú era mucho más efectivo que el resto de las numeraciones conocidas hasta el momento.

Un matemático persa llamado Omar Khayyam (1048-1131) descubrió métodos para resolver raíces cuadradas, cúbicas y de cualquier índice gracias a estas numeraciones. El más conocido de los matemáticos árabes es Mohammed

Ibn Musa Al-Khwarizmi (780-850) revolucionando el álgebra y sus métodos de cálculo. Se continuaron las investigaciones de Arquímedes acerca de las áreas y de los volúmenes y también se evolucionaron los problemas de óptica. Habas al-Hasib (792-870) y Nasir ad-Din at-Tusi (1201-1274) investigaron y crearon diversos tipos de trigonometrías planas y esféricas utilizando la función seno de los indios y el famoso teorema de Menelao de Alejandría (70-140).

Durante la Edad Media los trabajos de los árabes fueron los más reconocidos y tuvieron muchísima importancia junto con los descubrimientos anteriores de Grecia. Leonardo de Pisa, Leonardo Pisano o Leonardo Bigollo (1170-1250), también llamado Fibonacci posteriormente se basaría junto con Fray Luca Bartolomeo de Pacioli (1445-1517) en las matemáticas de los árabes para realizar sus estudios. Esto marcó una nueva época como es el Renacimiento, desarrollando sobre todo en Europa una gran cantidad de descubrimientos.

1.8. EJERCICIOS PROPUESTOS

1. **El viejo y el río**
   a) Simplifique la representación del problema del viejo y el río que viene en los apuntes
   b) Suponga que no hubiera el problema de que la oveja se come la alfalfa ¿cuántas veces tiene que cruzar en este caso?.
   c) ¿Hay otras soluciones al problema? ¿Cree que haya otra más corta?

Invente otro problema que tenga la misma solución abstracta que el del viejo y el río.

2. **Versión modificada del viejo y el río**
   Las condiciones del problema son las mismas solo que ahora en medio del río existe una isla. Si de un extremo al otro del río hay una distancia doble que de cualquier extremo a la isla ¿qué traslados tendrá que realizar el hombre para llevar todas sus pertenencias al otro extremo del río, de tal forma que tengan que recorrer la mínima distancia sin que el lobo se coma a la oveja y ésta última la alfalfa?

3. **Los esposos celosos**
   Tres hermosas desposadas con sus celosos maridos se encuentran a la orilla de un río y deciden cruzarlo. El pequeño bote que deben tomar para efectuar

el cruce solo tiene cabida para 2 personas. Para evitar cualquier situación comprometedora debe disponerse las travesías de tal manera que no se deje a ninguna mujer con un hombre, (en ningún momento) a menos que su esposo esté presente.

Construya un modelo simbólico a través del cual indique como deben cruzar el río y conteste las siguientes preguntas:

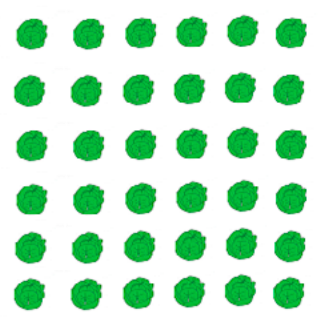
a) ¿Cuántas veces se cruza el río?
b) Si solo fueran dos parejas, ¿cuántas veces se cruzaría el río?
c) Si fueran cuatro parejas ¿podrían atravesar el río?

4. **Reparto de cerveza**
   Tres comerciantes deben repartir entre sí 21 barricas, de las cuales 7 están totalmente llenas de cerveza; 7 están llenas hasta la mitad y 7 están vacías. Se pregunta, como dividir estas barricas de tal forma, que a cada comerciante le corresponda la misma cantidad de cerveza y la misma cantidad de barricas, sin pasar cerveza de una barrica a otra.

5. **Repartiendo vino**
   Dos amigos tienen una jarra de vino de 8 litros desean repartírselo en partes iguales. Disponen además de la jarra de dos vasos vacíos, uno con capacidad de 5 litros y otro de 3 litros. ¿Cómo podrán dividirse el vino en partes iguales si los vasos no tienen graduación?

6. **Juan y sus hermanos**
   Juan tiene tres hermanitos; Pepito, Luisito y Jaimito; Juan tiene una moto en la cual debe llevarlos de la escuela a su casa, pero solo puede llevar uno a la vez. Sabiendo cómo es Pepito, Juan no quiere dejarlo solo con ninguno de sus otros hermanos, porque es capaz de pegarles si él no está presente para cuidarlos. ¿Cómo puede llevar Juan a sus hermanos a casa sin que se peleen? ¿Existe algún parecido entre el modelo utilizado en la resolución de este problema y el modelo empleado en el problema del viajero?

7. **Una elección ordenada**
   La figura siguiente representa el pequeño huerto de un laborioso aldeano en el que, como se ve, tenía éste plantadas 36 coles, formando cuadro con seis plantas en cada fila y seis en cada columna.



Apareció entonces un comprador que deseaba llevarse seis coles, y el aldeano las quiso escoger de tal forma que en el cuadro resultante siguiese habiendo un número par de plantas en todas las filas y también un número par en las columnas y las diagonales. ¿Qué plantas debió elegir para conseguir su propósito?

8. **Una regla mnemotécnica sobre π**
   El número π, razón de la circunferencia al diámetro, tan familiar a todos los estudiantes, ha sido calculado por muchos autores, cada vez con más decimales. Mucho antes de disponer de la ayuda de las computadoras, el matemático inglés W. Shanks (1812-1883) lo había calculado, en 1873, nada menos que con 707 cifras, y este valor figura grabado a lo largo del friso circular en que se apoya la cúpula del "Palais de la Decouverte" de París, si bien, desgraciadamente, las 150 últimas cifras decimales son erróneas. Para ninguna aplicación práctica son necesarias tantas cifras, bastando usualmente los valores aproximados 3, 14 o 3,1416 o 22/7. De todos modos, como regla mnemotécnica para recordar las primeras 32 cifras, se puede acudir a los siguientes versos (originales del ingeniero R. Nieto París, de Colombia):

> Soy π lema y razón ingeniosa
> de hombre sabio, que serie precisa
> valorando, enunció magistral.
> Por su ley singular, bien medido
> el grande orbe por fin reducido
> fue al sistema ordinario usual.

Se pide al estudiante que diga cuál es la regla mnemotécnica que se siguió.

9. **La planta maravillosa**
Sobre la superficie de un lago se encuentra una hoja de una planta acuática que cada día duplica su superficie. En ocho días se cubre así toda la superficie del agua ¿Cuántos días habría tardado en cubrirse esta misma superficie si el primer día en vez de una sola hoja se hubiesen puesto en el lago dos hojas iguales?

Capítulo 2
MATRICES

2.1. INTRODUCCIÓN

El problema central del Álgebra Lineal es la solución de ecuaciones lineales simultáneas; el caso más importante y el más simple, es aquél en que el número de incógnitas es igual al número de ecuaciones. Comenzaremos con dos ecuaciones lineales con dos incógnitas x_1 y x_2:

$$a_{11} x_1 + a_{12} x_2 = b_1$$
$$a_{21} x_1 + a_{22} x_2 = b_2$$

Donde a_{11}, a_{12}, a_{21}, a_{22}, b_1 y b_2 son números dados. Cada una de estas ecuaciones es la ecuación de una recta (en el plano x_1 x_2). La pendiente de la primera recta es: $-a_{11}/a_{12}$; y la pendiente de la segunda recta es: $-a_{21}/a_{22}$ (si $a_{12} \neq 0$ y $a_{22} \neq 0$). Una solución de este sistema es una pareja de números (x_1 , x_2) que satisfacen el sistema.

Otra interpretación consiste en observar el sistema en forma de columnas, se tiene entonces:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{11} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Y aquí se trata de encontrar la combinación de los vectores en el lado izquierdo que dé como resultado el lado derecho.

Recordando la definición de vector, se tiene lo siguiente:

Definición 2.1
Un vector es todo segmento de recta dirigido en el espacio. Cada vector posee unas características que son:

» *Origen:* También denominado punto de aplicación. Es el punto exacto sobre el que actúa el vector.
» *Módulo:* Es la longitud o tamaño del vector. Para hallarla es preciso conocer el origen y el extremo del vector, pues para saber cuál es el módulo del vector, debemos medir desde su origen hasta su extremo.
» *Dirección:* Viene dada por la orientación en el espacio de la recta que lo contiene.
» *Sentido:* Se indica mediante una punta de flecha situada en el extremo del vector, indicando hacia qué lado de la línea de acción se dirige el vector. El sistema de referencia que usaremos, como norma general, es el Sistema de Coordenadas Cartesianas.

Volviendo al tema de la solución de un sistema de ecuaciones en el ejemplo siguiente se ilustran estas dos interpretaciones.

Ejemplo 2.1
Resuelva el siguiente sistema de ecuaciones:

$$2x_1 + 4x_2 = 2 \quad (1)$$
$$4x_1 + 11x_2 = 1 \quad (2)$$

Multiplicando la primera ecuación por dos y restándola a la segunda se tiene:

$$4x_1 + 8x_2 = 4 \quad (1)$$
$$4x_1 + 11x_2 = 1 \quad (2)$$

De donde $x_2 = -1$ y $x_1 = 3$, entonces la solución es $(x_1, x_2) = (3, -1)$, y geométricamente se representa en la figura 2.1.

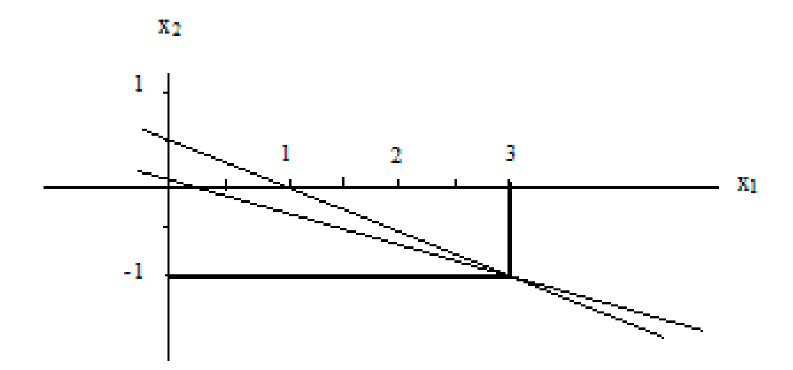


Figura 2.1 solución gráfica de un sistema de dos ecuaciones

Si seguimos la segunda interpretación geométrica, consideramos entonces las columnas del sistema y no los renglones, y se tiene:

$$x_1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 11 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Como ya decíamos, se quiere encontrar la combinación de los vectores en el lado izquierdo que dé como resultado el lado derecho. Geométricamente se representa en la figura 2.2:

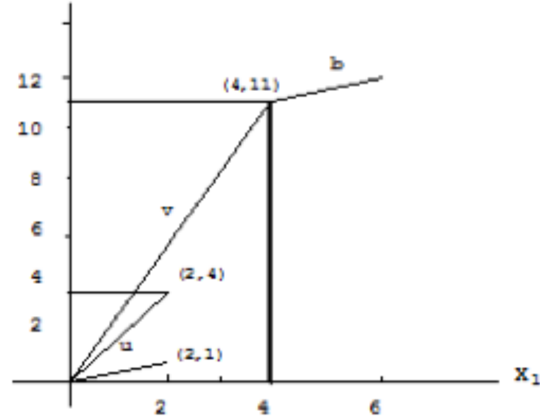


Figura 2.2 solución geométrica de un sistema de vectores

Multiplicamos el vector $[2,4]^t$ por 3 y su longitud se triplica:

$$3 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

Cuando el vector se multiplica por 1 su dirección se invierte:

$$- \begin{bmatrix} 4 \\ 11 \end{bmatrix} = \begin{bmatrix} -4 \\ -11 \end{bmatrix}$$

Geométricamente, al sumar vectores se coloca un vector al inicio de donde el otro termina, o algebraicamente sumando sus componentes correspondientes:

$$3\begin{bmatrix}2\\4\end{bmatrix}-1\begin{bmatrix}4\\11\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}$$

Con ambos métodos se obtiene el mismo resultado $x_1 = 3$ y $x_2 = -1$

Podemos plantearnos ahora las siguientes preguntas:

» ¿Cuándo tiene solución el sistema?
» ¿Cuántas soluciones tiene?

Veamos algunos ejemplos:

Ejemplo 2.2
Considere el sistema:

$$x_1 - x_2 = 7$$
$$x_2 + x_2 = 5$$

Si sumamos las dos ecuaciones obtenemos: x_1=6 y x_2=1. Esta es la solución del sistema, es decir, cumple con las dos ecuaciones y además es única, lo que se puede comprobar geométricamente:

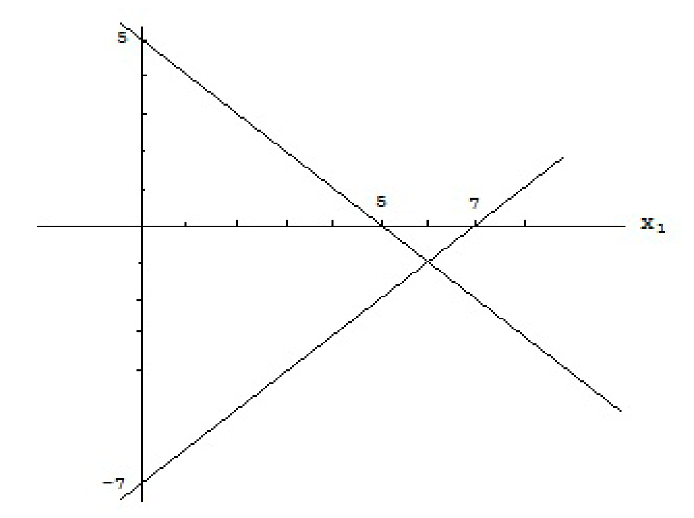


Figura 2.3
solución única de un sistema de dos ecuaciones con dos incógnitas

En la figura 2.3 se puede ver que la solución es el punto de intersección de las dos rectas.

Ejemplo 2.3

Considere el sistema:

$$x_1 - x_2 = 7 \quad (1)$$
$$2x_2 - 2x_2 = 14 \quad (2)$$

Se puede ver claramente que la ecuación (2) es resultado de multiplicar la primera por 2, de aquí tenemos que $x_2 = x_1 - 7$.

Así, el par $(x_1, x_2 - 7)$ es una solución al sistema para cualquier número real x_1 por lo tanto el sistema tiene un número infinito de soluciones, tantas como valores pueda tomar x_1. Por ejemplo, los siguientes pares son soluciones:

(7,0), (0,7), (8,1), (1,6), (3,4), y (2,9)

Lo podemos verificar geométricamente en la figura 2.4:

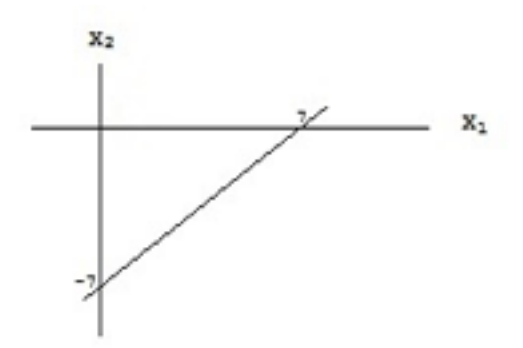


Figura 2.4 Infinidad de soluciones para un sistema de dos ecuaciones con dos incógnitas

Ejemplo 2.4

Considere el sistema:

$$x_1 - x_2 = 7 \quad (1)$$
$$2x_1 - 2x_2 = 13 \quad (2)$$

Si multiplicamos la primera ecuación por dos, tenemos:

$$2x_1 - 2x_2 = 14$$

Contrariamente a lo que se afirma en la segunda ecuación; por lo tanto, el sistema no tiene solución, como podemos ver en la figura 2.5:

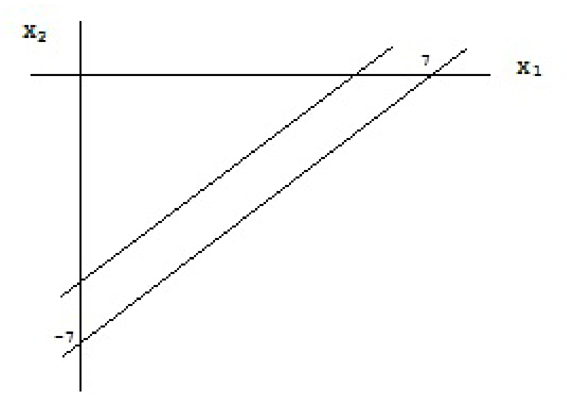


Figura 2.5 Sistema de dos ecuaciones con dos incógnitas sin solución

Considere nuevamente el sistema de dos ecuaciones con dos incógnitas:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \qquad (1)$$

Si multiplicamos la primera ecuación por a_{22} y la segunda por a_{12} tenemos:

$$\begin{aligned} a_{11}a_{22}x_1 + a_{12}a_{22}x_2 &= a_{22}b_1 \\ a_{12}a_{21}x_1 + a_{12}a_{22}x_2 &= a_{12}b_2 \end{aligned} \qquad (2)$$

El sistema (1) y el sistema (2) son equivalentes. Esto significa que cualquier solución del sistema (1) es una solución del sistema (2) y viceversa. Restando la segunda ecuación de la primera en el sistema (2) se tiene:

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2$$

Si ($a_{11}\ a_{22}\ a_{12}\ a_{21}$) $\neq 0$, se puede dividir entre esta cantidad y se tiene:

$$x_1 = \frac{a_{22}\,b_1 - a_{12}\,b_2}{a_{11}\,a_{22} - a_{12}\,a_{21}}$$

Entonces se sustituye este valor en el sistema (1) y se procede a resolver para x_2, encontrándose así la única solución del sistema.

> ***Definición 2.2***
> Se define el DETERMINANTE DEL SISTEMA (1) como:
>
> Determinante del sistema (1) = $(a_{11}\ a_{22}\ a_{12}\ a_{21})$

Y se ha mostrado que si el determinante del sistema (1) es diferente de cero, entonces el sistema tiene una única solución. Si es igual a cero, no tiene solución o tiene un número infinito de soluciones.

2.2. MATRICES

Como veremos más adelante, la solución de un sistema de ecuaciones lineales se trata eficientemente simplificándolo en un arreglo que prescinde de las letras y los signos de igual. El resultado es un ordenamiento de números conocido como matriz.

> ***Definición 2.3***
> Una matriz es un arreglo de datos ordenados en renglones y columnas, normalmente entre corchetes o paréntesis. A una matriz con m renglones y n columnas se le denomina matriz de m por n.

Las matrices son una herramienta fundamental para realizar cálculos eficientes en álgebra lineal.

Por ejemplo:

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Es la matriz de los coeficientes del sistema de ecuaciones lineales:

$$x_1 - x_2 = 7$$
$$x_1 + x_2 = 5$$

En general, A es una matriz de $m \times n$ si:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Si $n=m$, la matriz es cuadrada, y si $n \neq m$, la matriz es rectangular.

OPERACIONES CON MATRICES

> ***Definición 2.4***
> *SUMA ALGEBRÁICA DE MATRICES*. Sean A y B dos matrices de orden $m \times n$, la suma o diferencia de ambas $A \pm B$ es otra matriz C, de orden $m \times n$, en la que cada elemento de C es igual a la suma o diferencia de los elementos correspondientes de A y B. Por ejemplo:

Ejemplo 2.5

Sean A y B como sigue:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 3 & 0 \\ -1 & 2 & 5 \end{bmatrix}$$

Entonces $A+B$ está dada por:

$$A+B = \begin{bmatrix} 1+2 & 2+3 & 3+0 \\ 0-1 & 1+2 & 4+5 \end{bmatrix} \quad A+B = \begin{bmatrix} 3 & 5 & 3 \\ -1 & 3 & 9 \end{bmatrix}$$

Dos matrices de distinto orden no se pueden sumar ni restar. La suma de matrices A es otra matriz del mismo orden que A cuyos elementos son iguales a los correspondientes de A multiplicados por k.

Ejemplo 2.6
Considere la matriz A como sigue:

Se tiene:

$$A=\begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}$$

En particular, $-A$ se conoce como la **matriz opuesta** de A.

$$A+A+A=\begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}+\begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}+\begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}=\begin{bmatrix} 3 & -6 \\ 6 & 9 \end{bmatrix}=3A$$

Multiplicación de matrices

Considere el sistema lineal:

$$\begin{aligned} 2x_1-3x_2 &= 7 \\ 3x_1-4x_2 &= 2 \end{aligned} \quad (3)$$

Los datos importantes para este sistema lineal están contenidos en la matriz de coeficientes:

$$A=\begin{bmatrix} 2 & -3 \\ 3 & -4 \end{bmatrix}$$

Y el vector columna:

$$A=\begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

De las constantes que aparecen en el lado derecho.

Sea también:

$$x=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

CONTENIDO

El vector columna de las incógnitas del sistema (3). Es costumbre abreviar el sistema (3) como:

$Ax=b$

O bien:

$$\begin{bmatrix} 2 & -3 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

Esta multiplicación sugiere la multiplicación de la matriz A por un vector columna x. Se define entonces:

Definición 2.5
PRODUCTO PUNTO O ESCALAR. Dados dos vectores a y b con el mismo número de componentes, el producto punto o escalar de a y b está dado por:

$a{\cdot}b = a_1\, b_1 + a_2\, b_2 + \cdots + a_n b_n$

Es esencial diferenciar entre los conceptos de vector y número. El término **escalar** se suele utilizar en álgebra lineal en lugar de número. El producto punto también se llama **producto escalar**, pues a·b es una cantidad escalar.

Ejemplo 2.7
Efectúe el producto escalar de los vectores a y b donde:

a) $a = (1, -3, 4)$ y $b = (2, 0, 8)^t$

Se tiene entonces: $a \cdot b = (1)(2) + (-3)(0) + (4)(8) = 2 + 0 + 32 = 34$

b) $a = (1, 5, 6)$ y $b = (2, 4)^t$

Se observa que $a{\cdot}b$ no está definido, ya que no tienen el mismo número de componentes.

El producto AB de dos matrices A y B se forma tomando todos los productos punto posibles de los vectores renglón de A por los vectores columna de B.

Para que estos productos estén definidos, el número de componentes de cada renglón de A debe ser igual al número de componentes de cada columna de B. Si AB está definido y A es de $m \times n$, de modo que sus renglones tengan n componentes, entonces B debe ser de $n \times s$ para que las columnas de B tengan también n componentes.

$$AB = (c_{ij}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & b_{1s} \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nj} & b_{ns} \end{bmatrix}$$

En términos de sumatoria tenemos:

$$c_{ij} = a_{i1}\, b_{1j} + \cdots + a_{in}\, b_{nj} = \sum_{k=1}^{n} a_{ik}\, b_{kj}$$

AB está definida solo cuando el número de columnas de A es igual al número de renglones de B. La matriz producto tiene tamaño:

(Número de renglones de A) × (Número de columnas de B)

Ejemplo 2.8

Efectúe el producto de las siguientes matrices A y B:

$$A = \begin{bmatrix} -2 & 3 & 2 \\ 4 & 6 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & -1 & 2 & 5 \\ 3 & 0 & 1 & 1 \\ -2 & 3 & 5 & 3 \end{bmatrix}$$

$$AB = \begin{bmatrix} -3 & 8 & 9 & -1 \\ 38 & -10 & 4 & 20 \end{bmatrix}$$

2.3 . PROPIEDADES DE LAS MATRICES

Las propiedades de las matrices nos permiten operar con ellas de tal manera que los resultados sean correctos y no se hagan suposiciones que conlleven a errores futuros.

Consideremos ahora las siguientes propiedades de las matrices, donde se involucran la suma y el producto de estas.

1. $A(B + C) = AB + AC$ (primer propiedad distributiva).
2. $(A + B)C = AC + BC$ (segunda propiedad distributiva).
3. $A(BC) = (AB)C$ (propiedad asociativa).

Sin embargo:

4. $AB \neq BA$ en general.
5. $AB = 0$ no implica necesariamente que A=0 ó B=0.
6. $AB = AC$ no implica necesariamente que $B = C$.

Si A y B son dos matrices cuadradas y se verifica que $AB = BA$, dichas matrices se llaman **permutables o conmutativas**.

Ejemplo 2.9

Sean las matrices A, B, C y D como sigue:

$$A=\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \quad B=\begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix} \quad C=\begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix} \quad D=\begin{bmatrix} 3 & 7 \\ 0 & 0 \end{bmatrix}$$

Se tiene entonces que $AB = AC$ y $B \neq C$:

$$AB = AC\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

Además: AD=0, con $A \neq 0$ *y* $D \neq 0$.

Casos particulares de matrices cuadradas

Las matrices que se consideran en esta parte tienen una característica especial al aplicarse en estadística, econometría y ecuaciones diferenciales por mencionar algunas áreas.

Definición 2.6
Una matriz A de manera que $A^{k+1}= A$, siendo k un número entero positivo, se llama **periódica**. Si k es el menor número entero positivo para el cual $A^{k+1}=A$, la matriz A tiene periodo k. Si $k = 1$, esto es: $A^2=A$, la matriz A se llama **idempotente**.

Ejemplo 2.10

a) Sea la matriz A como sigue:

$$A=\begin{bmatrix} 1 & -2 & -6 \\ -3 & 2 & 9 \\ 2 & 0 & -3 \end{bmatrix}$$

Es periódica de periodo 2, ya que $A^3 = A$.

b) La matriz A es idempotente, esto es: $A^2 = A$, con A definida como sigue:

$$A=\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$$

Definición 2.7
Una matriz A tal que $A^p=0$ siendo p un número entero positivo se llama **nilpotente**. Si p es el menor número entero positivo para el cual $A^p = 0$, la matriz A es nilpotente de índice p.

Ejemplo 2.11

Sea la matriz A como sigue:

$$A=\begin{bmatrix} 1 & -2 & -6 \\ -3 & 2 & 9 \\ 2 & 0 & -3 \end{bmatrix}$$

Es nilpotente de orden 3, $A^3 = 0$.

Definición 2.8

MATRIZ TRANSPUESTA. La matriz transpuesta de una matriz A de orden $m \times n$ es la matriz A^t de orden $n \times m$ que se obtiene permutando los renglones por las columnas.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$A^t = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

Ejemplo 2.12

Encuentre la transpuesta de la siguiente matriz A:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Observe que el elemento a_{ij} de A (renglón i, columna j) es el a_{ji} de A^t (renglón j, columna i). O bien, si

$$A = [a_1 \cdots a_n] \qquad A^t = \begin{bmatrix} a_1 \\ \cdots \\ a_n \end{bmatrix}$$

El producto escalar de dos vectores columna $<x, y>$ está dado por:

$$<x, y> = x^t\, y = x_1\, y_1 + x_2\, y_2 + \cdots + x_n y_n$$

O bien: $\quad <x, y> = y^t\, x = x_1\, y_1 + x_2\, y_2 + \cdots + x_n y_n$

De donde: $\quad x^t y = y^t x = \sum_{i=1} x_i y_i$

Propiedades de las matrices transpuestas

Sean A^t y B^t las transpuestas de las matrices A y B respectivamente, y k un escalar cualquiera; de aquí que:

a) $(A^t)^t = A$
b) $(kA)^t = kA^t$

Además, se verifica que si $A = A^t$, entonces la matriz A es simétrica.

2.4. RESUMEN

1. Una matriz de $m \times n$ es una disposición rectangular ordenada de números que contiene m filas y n columnas.

2. Una matriz de $mx1$es un vector columna con m componentes y una matriz de $1xn$ es un vector fila con n componentes.

3. El producto punto del vector a, con componentes $a_1, a_2, \cdots, a_n$, por el vector b, con componentes $b_1, b_2, \cdots, b_n$,es el escalar:

   $a \cdot b = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$

4. El producto AB de una matriz A de $m \times n$ y de una matriz B de nxs, es la matriz C de mxs cuyo elemento c_{ij}en la i-ésima fila y la j-ésima columna es el producto punto del i-ésimo vector fila de A por el j-ésimo vector columna de B. En general, $AB \neq BA$.

5. Si $A = (a_{ij})$ y $B = (b_{ij})$ son matrices del mismo tamaño, entonces $A+B$ es la matriz de ese tamaño con elemento $a_{ij}+b_{ij}$ en la i-ésima fila y la j-ésima columna.

6. Para cualquier matriz A y escalar r, la matriz rA se halla multiplicando cada elemento de A por r.

7. La transpuesta de una matriz A de $m \times n$ es la matriz A^t de $n \times m$ que tiene como k-ésimo vector fila al k-ésimo vector columna de A.

2.5. APLICACIONES DE MATRICES

En esta sección se presentan dos aplicaciones de matrices tanto de teoría de gráficas como de teoría de juegos, la lectura de ellos es opcional en este libro, sin embargo, se recomienda para tener una mayor idea de cómo se aplican las matrices en la vida real, existen muchas otras aplicaciones, sin embargo, por cuestiones de espacio nos enfocaremos solamente a estas dos.

2.5.1. TEORÍA DE GRÁFICAS

(Tomado de *Aplicaciones de álgebra lineal*, Rorres y Anton (1977))

En muchos problemas de la ingeniería en particular y en general de diferentes áreas del conocimiento se tienen casos de conjuntos que tienen un número finito de elementos que de alguna forma están relacionados entre sí.

Desde hace siglos los científicos han hecho uso de ellas por ejemplo Los físicos teóricos han utilizado las gráficas de múltiples maneras. Por ejemplo, en mecánica se utilizan gráficas en las que los vértices representan moléculas, y cuando dos de ellos son adyacentes indican algún tipo de interacción física, tal como repulsión o atracción magnética. De manera similar, Feynman usó gráficas en las que los vértices representan partículas físicas, y las aristas trayectorias de estas después de colisionar.

http://www.revistaciencias.unam.mx/es/86-revistas/revista-ciencias-67/747-un-vistazo-a-la-teoria-de-graficas.html.

Kirchhoff también hizo descubrimientos en la teoría de gráficas cuando intentaba resolver el sistema de ecuaciones lineales simultáneas obtenido por el paso de corriente en cada circuito de una red eléctrica. Así, abstrajo de la red eléctrica —con sus resistencias, condensadores, inductancias y demás—, una estructura consistente solo de puntos y líneas sin indicar el tipo de elemento eléctrico que representaban. Mediante la "gráfica subyacente" a la red eléctrica, notó que el problema original podía ser resuelto más fácilmente sin considerar los ciclos que se formaban, utilizando lo que hoy conocemos como "árbol generador".

http://www.revistaciencias.unam.mx/es/86-revistas/revista-ciencias-67/747-un-vistazo-a-la-teoria-de-graficas.html.

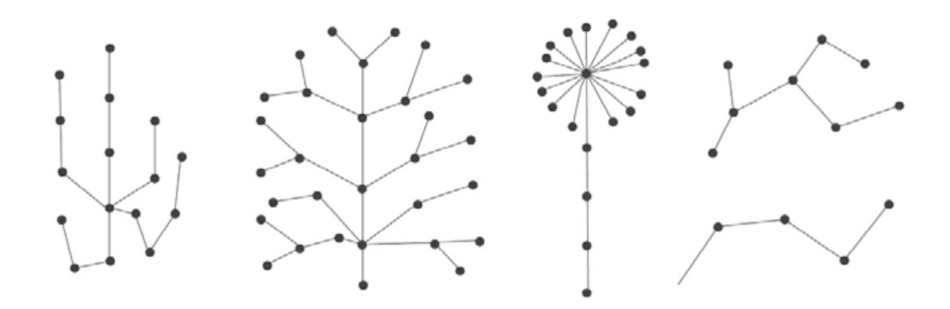


Figura 2.6 Árbol generador Fuente http://www.revistaciencias.unam.mx/es/86-revistas/revista-ciencias-67/747-un-vistazo-a-la-teoria-de-graficas.html

Las gráficas se utilizan al diseñar circuitos integrados impresos en chips de silicón, que son usados en dispositivos electrónicos —ésta es una de las aplicaciones más importantes de lo que conocemos como gráficas planas— que deben ser diseñados de tal modo que las porciones conductoras no se crucen entre sí.

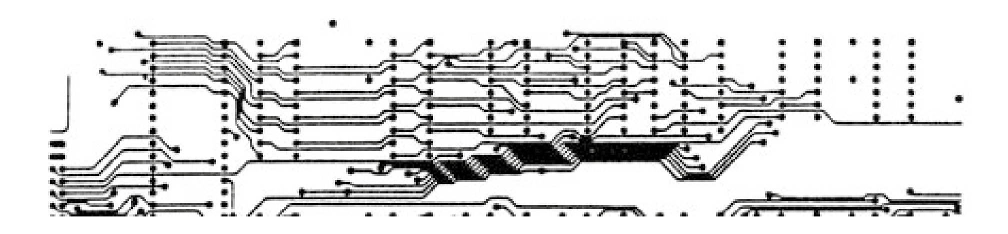


Figura 2.7 Circuito integrado. Fuente http://www.revistaciencias.unam.mx/es/86-revistas/revista-ciencias-67/747-un-vistazo-a-la-teoria-de-graficas.html

En investigación de operaciones, las gráficas juegan un papel relevante —ahí se les conoce como redes y casi siempre son dirigidas, es decir, a cada arista se le agrega "una flecha" que indica una dirección— y se utilizan principalmente para la planeación de actividades y el diseño de sistemas de comunicación o distribución de bienes y servicios. Por ejemplo, se pueden considerar conjuntos de personas en una empresa, de animales en una granja, de países a nivel político y económico, de empresas, equipos deportivos o de ciudades. La relación entre dos miembros de un conjunto, *A* y *B*, pueden ser, que la persona *A* domine a la persona *B*, que el animal *A* se alimente del animal *B*, que el país *A* apoye

militarmente al país *B*, que la compañía *A* venda su producto a la compañía *B*, que el equipo deportivo *A* gane sistemáticamente al equipo *B*, o que la ciudad *A* tenga un vuelo directo a la ciudad *B*.

Las gráficas están relacionadas con las redes en su estructura y diseño, aunque tienen objetivos diferentes, en las redes se tiene como objetivo encontrar el flujo en la red y en la gráfica el objetivo es una relación de tipo estructural, en particular se estudiarán gráficas dirigidas para lo cual se tiene la siguiente definición.

Definición 2.5.1 Gráficas Dirigidas

Una gráfica dirigida es un conjunto finito de elementos, $P_1, P_2, \cdots, P_n$, asociado a una colección finita de pares ordenados, (P_i, P_j) formados por dos elementos distintos del conjunto y sin que se repita ningún par. A los elementos del conjunto se les llama *vértices* y a los pares ordenados, *líneas dirigidas*. La notación $P_i \rightarrow P_j$ (*léase P_i se relaciona con P_j*) indicará que la línea dirigida la (P_i, P_j) pertenece a la gráfica. Se obtiene un diagrama de una gráfica dirigida, (Fig. 2.8), representando a los vértices como puntos de un plano y a $P_i \rightarrow P_j$ por una línea, recta o curva, trazada del vértice P_i al vértice P_j, con una flecha que apunte de P_i a P_j. Si las expresiones $P_i \rightarrow P_j$ y $P_j \rightarrow P_i$ son simultáneamente válidas (en cuyo caso se puede escribir $P_i \leftrightarrow P_j$), se traza una sola línea entre (P_i y P_j) con dos flechas de sentidos contrarios (como entre P_2 y P_3) en la figura).

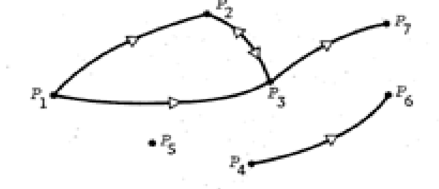


Figura 2.8 Diagrama de una gráfica dirigida

Observe, que como ocurre en la figura 2.8, una gráfica dirigida puede tener "componentes" separadas debidas a grupos de vértices que solo se unen entre ellos y que algunos vértices, como pueden quedar aislados. Además, como no se permite la expresión , un vértice no queda unido consigo mismo más que a través de algún otro vértice.

En la figura 2.9 aparecen tres diagramas que corresponden a otros ejemplos de gráficas dirigidas.

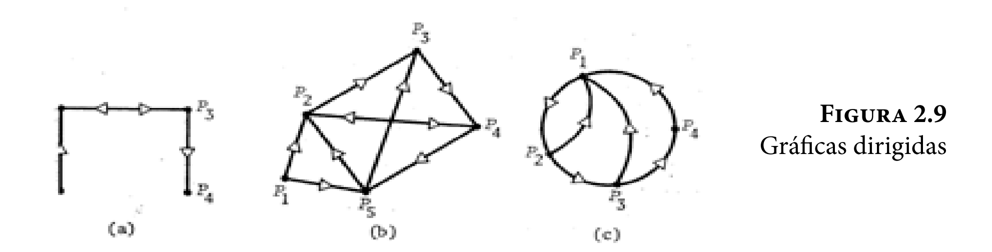


FIGURA 2.9
Gráficas dirigidas

A una gráfica dirigida de *n* vértices, se le puede asociar una matriz de *n*, $M=[m_{ij}]$, que recibe el nombre de matriz de los vértices, en algunos libros se conoce como matriz de adyacencia. Sus elementos se definen por medio de

1, *si* $P_i \rightarrow P_j$
$m_{ij} = 0$, *en los demás casos*

Para $i, j = 1, 2, \cdots n$.A las gráficas de la figura 2.9, les corresponden las siguientes matrices de los vértices:

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad M = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix} \quad M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Figura 2.9(a) Figura 2.9(b) Figura 2.9(c)

Por definición, las matrices de los vértices tienen las dos propiedades siguientes:
(i) Todos los valores de entrada son 0 ó 1.
(ii) Todos los valores de entrada sobre la diagonal principal son 0.

Ejemplo 2.13

Una cierta familia está formada por la madre, el padre, una hija y dos hijos. La influencia o el poder que unos de sus miembros ejercen sobre los otros se describen a continuación: la madre sobre la hija y el hijo mayor, el padre sobre los dos hijos, la hija sobre el padre, el hijo mayor sobre el hijo menor y el hijo menor sobre la madre. Un modelo del patrón de influencia en esta familia se obtiene con una gráfica dirigida cuyos vértices son los cinco miembros de la familia. Si el miembro A ejerce influencia sobre el miembro *B*, se escribirá $A \rightarrow B$.

En la figura 2.10 aparece la gráfica dirigida resultante, donde los diferentes miembros de la familia han sido llamados de la siguiente forma: madre por M,

padre por F, hija por D, hijo mayor por OS e hijo menor por YS. La matriz de los vértices correspondiente es:

$$\begin{array}{c} \\ M \\ F \\ D \\ OS \\ YS \end{array} \begin{array}{c} \begin{array}{ccccc} M & F & D & OS & YS \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

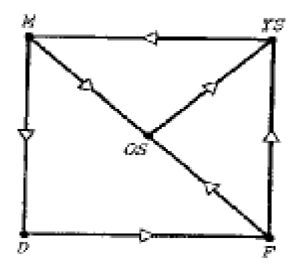


Figura 2.10 Gráfica dirigida del ejemplo 2.13

En el ejemplo 2.13, el padre no tiene influencia directa sobre la madre, es decir, $F \to M$ no es cierta. Sin embargo, tiene influencia sobre su hijo menor, quien, a su vez, la tiene sobre la madre. Esto se escribe, $F \to YS \to M$, y se dice que haga una *liga de dos pasos* entre F y M. En forma análoga, $M \to D$ es una *liga de un paso* y $F \to OS \to YS \to M$ es una *liga de tres pasos*. Ahora, se verá la forma de determinar el número de ligas de r pasos (r=1, 2, ⋯) que hay entre un vértice P_i y otro P_j de una gráfica dirigida cualquiera. (Se incluirá el caso en que P_i y P_j coinciden.) El número de ligas de 1 paso entre P_i y P_j es, simplemente m_{ij}. Es decir, puede haber ninguna o una liga de un paso de P_i a P_j , dependiendo de que m_{ij} sea cero o uno. El número de ligas de dos pasos se obtiene con el cuadrado de la matriz de los vértices. Si $m_{ij}^{(2)}$ es el elemento de orden de (i, j) de M^2, se tiene:

$$m_{ij}^{(2)} = m_{i1}\, m_{1j} + m_{i2}\, m_{2j} + \cdots + m_{in}\, m_{nj} \qquad (2.5.1)$$

Ahora, si $m_{i1} = m_{1j} = 1$, hay una liga de dos pasos, $P_i \to P_1 \to P_j$, entre P_i y P_j.

Pero si una de las dos, m_{i1} ó m_{1j}, es cero, esta liga no se da. En consecuencia, $P_i \to P_1 \to P_j$ es una liga de dos pasos si, y solo si, $m_{i1}\, m_{1j} = 1$. De forma semejante, para cualquier valor k= 1, 2, ⋯, n, $P_i \to P_k \to P_j$ es una liga de dos pasos entre P_i y P_j si, y solo si, el término $m_{ik}\, m_{kj}$ del segundo miembro de (2.5.1) vale uno; en caso contrario, el término es cero.

En consecuencia, el segundo miembro de (2.5.1) da el total de ligas de dos pasos que hay entre P_i y P_j.

El mismo argumento se utiliza para obtener el número de ligas de 3 pasos, de 4 pasos, etc., que hay entre P_i y P_j. En general, se puede decir que:

Teorema 2.1

Si M es la matriz de los vértices de una gráfica dirigida y $m_{i1}^{(r)}$ es el elemento de orden (i,j) de M^r, $m_{i1}^{(r)}$ es el número de ligas de r pasos que hay entre P_i y P_j.

Ejemplo 2.14

La figura 2.11 es el mapa de las rutas de una pequeña línea aérea que da servicio a cuatro ciudades P_1, P_2, P_3, y P_4. Como es una gráfica dirigida, la matriz de los vértices será:

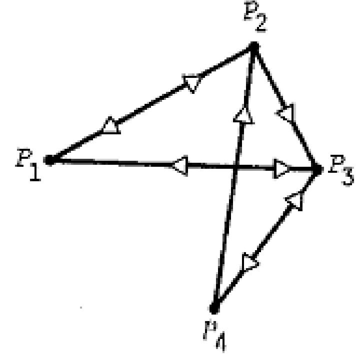


FIGURA 2.11 Gráfica de las rutas de una línea aérea

La matriz de vértices asociada es:

$$M = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Además, las matrices M^2 y M^3son:

$$M^2 = \begin{bmatrix} 2 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix} \qquad M^3 = \begin{bmatrix} 1 & 3 & 3 & 1 \\ 2 & 2 & 3 & 1 \\ 4 & 0 & 2 & 2 \\ 1 & 3 & 3 & 1 \end{bmatrix}$$

Si se quiere saber cuántas conexiones hay entre las ciudades P_4 y P_3, lo que se puede determinar mediante el teorema 2.1. Como $m_{43} = 1$, hay una conexión o liga de 1 paso; como $m_{43}^{(2)} = 1$, hay una conexión de 2 pasos; y como $m_{43}^{(3)} = 3$, hay tres conexiones de 3 pasos. Se puede verificar en la figura 2.9.

conexiones de 1 *paso de* P_4 *a* P_3:	$P_4 \rightarrow P_3$
conexiones de 2 *paso de* P_4 *a* P_3:	$P_4 \rightarrow P_2 \rightarrow P_3$
conexiones de 3 *paso de* P_4 *a* P_3:	$P_4 \rightarrow P_3 \rightarrow P_4 \rightarrow P_3$
	$P_4 \rightarrow P_2 \rightarrow P_1 \rightarrow P_3$
	$P_4 \rightarrow P_3 \rightarrow P_1 \rightarrow P_3$

Otra relación interesante en la teoría de gráficas es una asociación que se define de la siguiente manera:

> ***Definición 2.5.2 Asociación***
> Un subconjunto de una gráfica dirigida recibe el nombre de asociación si satisface las siguientes condiciones:
> (i) El subconjunto contiene por lo menos tres vértices.
> (ii) Para cada par de vértices P_i y P_j del subconjunto, son ciertas las expresiones $P_i \to P_j$ *y* $P_j \to P_i$.
> (iii) El subconjunto es lo más grande posible, es decir, no se puede agregar otro vértice sin dejar de satisfacer la condición (ii).

Esta definición sugiere que las asociaciones son subconjuntos máximos dentro de las cuales existe una comunicación perfecta. Por ejemplo, si los vértices representan ciudades y $P_i \to P_j$ significa que hay un vuelo directo entre P_i y P_j, entonces hay un vuelo directo entre dos ciudades cualesquiera, de la asociación y en cualquiera de las dos direcciones.

Ejemplo 2.15

La gráfica dirigida que aparece en la figura 2.12 (que podría representar el mapa de las rutas de una línea aérea) tiene dos asociaciones:

$\{P_1, P_2, P_3, P_4\}$ y $\{P_3, P_4, P_6\}$

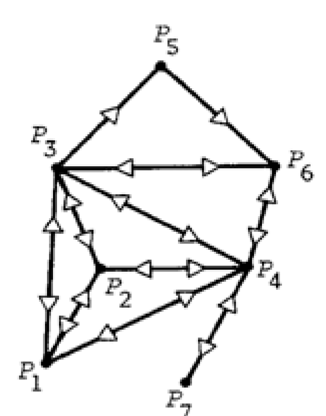


Figura 2.12 Rutas de una línea aérea

Es importante aquí hacer otra definición muy útil en teoría de gráficas.

Definición 2.5.3 Clique
Es cualquier subgráfica de $G=(V,E)$ isomórfa a la gráfica completa K_i, donde $1 \le i \le n$ y $n=|V|$. Esto significa que siempre podemos particionar los vértices de la gráfica en cliques.

Ejemplo 2.16

La gráfica dirigida que aparece en la figura 2.10 (que podría representar el mapa de las rutas de una línea aérea) tiene dos asociaciones:

$$\{P_1, P_2, P_3, P_4\} \text{ y } \{P_3, P_4, P_6\}$$

En este ejemplo se observa que en una gráfica dirigida puede haber más de un clique y que un vértice puede pertenecer simultáneamente a dos o más de ellos. Como en la figura 2.12.

En las gráficas dirigidas simples, los cliques pueden encontrarse por inspección, pero para los casos complicados conviene tener un procedimiento que permita detectarlos. A continuación, se analizará, un teorema que identifica a los vértices que pertenecen a un clique y para ello se definirá una matriz $S=[sij]$ que está relacionada con una gráfica dirigida de la siguiente forma:

$$Sij = \begin{cases} 1, \text{ si } Pi \leftrightarrow Pj \\ 0, \textit{ en los demás casos} \end{cases}$$

La matriz S determina una gráfica dirigida que se diferencia de la original en que han desaparecido o se han eliminado, las líneas dirigidas que solo tienen una flecha. Por ejemplo, si la gráfica dirigida original es la que se ilustra en la figura 2.13(a), la gráfica dirigida que tiene a S como la matriz de los vértices es la que aparece en la figura 2.13(b). Como alternativa, S se puede obtener a partir de la matriz de los vértices, M de la gráfica original, haciendo que $sij=1$ cuando $mij=mji=1$ y en caso contrario, $sij=0$.

Usando la matriz S el teorema antes mencionado se enuncia de la siguiente forma:

Teorema 2.2

Si S^3_{ij} es el elemento de orden (i,j) de S^3 de, un vértice P_i pertenece a algún clique si, y solamente si, $s_{ii} \neq 0$

Ejemplo 2.17
Suponga que una gráfica dirigida tiene por matriz de los vértices:

$$M = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Entonces

$$S = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad S^3 = \begin{bmatrix} 0 & 3 & 0 & 2 \\ 3 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & 1 & 0 \end{bmatrix}$$

Como todas las entradas a la diagonal de S^3 son cero, por el teorema 2.2 puede decir que la gráfica no tiene cliques.

Gráficas de dominación

Hay muchos grupos de individuos o de animales en los que existe un orden bien definido del sentido de la dominación entre dos miembros cualesquiera. Dados dos individuos cualesquiera, A y B, sucede que A domina a B o bien que B domina a A, pero no ambas. En términos de una gráfica dirigida, en la que $P_i \rightarrow P_j$ significa que P_i domina a P_j para todas las parejas diferentes sucede que $P_i \rightarrow P_j$ o que $P_i \leftarrow P_j$, pero nunca $P_i \leftrightarrow P_j$. En general, se puede dar la siguiente definición:

> ***Definición 2.5.4*** **Una gráfica de dominación**
> Es una gráfica dirigida de tal forma que, para cualquier par de vértices, P_i, *y* P_j se verifica que $P_i \rightarrow P_j$ o que $P_j \rightarrow P_i$, pero no las dos.

Un ejemplo de una gráfica dirigida que cumple con esta definición es la que corresponde a una liga de n equipos deportivos en la que cada equipo juega solo una vez contra todos los demás y donde no se permiten los empates. Si $P_i \rightarrow P_j$ significa que el equipo José Antonio Sánchez Calderón o P_i le gana al equipo P_j

en la única vez que se encuentran durante el torneo, de inmediato se ve la definición 5.2.4 si se satisface. Por esta razón, a las gráficas de dominación, se les suele llamar *torneos*.

En la figura 2.13 se muestran tres gráficas de dominación de tres, cuatro y cinco vértices. En ellas, los vértices sombreados tienen una propiedad muy interesante: cada uno de ellos se une, por medio de una liga de 1 o 2 pasos, con cualquiera de los otros vértices de la gráfica. En cierto sentido, estos vértices son más "potentes" que los que no tienen esta propiedad. En un torneo deportivo, un vértice de tal naturaleza corresponde a un equipo*A* que le gana a otro equipo*B*, además, a un equipo que le gana a *B*.

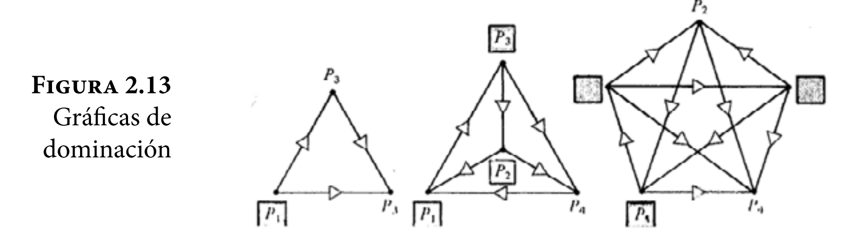


Figura 2.13 Gráficas de dominación

2.5.2. TEORÍA DE JUEGOS

(Tomado de *Aplicaciones de Álgebra Lineal*, Grossman (1988))

Juegos entre dos personas: Estrategias puras

La teoría moderna de juegos fue desarrollada en la década de 1940 por John von Neumann y Oskar Morgenstern (1944), para dar un marco matemático general a la economía. Las ideas principales de esta teoría se extrajeron de juegos bien conocidos como el ajedrez, el bridge, el solitario, dominó y damas. La teoría general se desarrolló sin hacer referencia concreta a ningún juego en particular. La teoría de juegos se puede aplicar al análisis de cualquier comportamiento competitivo, incluyendo los juegos ordinarios, la economía, la guerra y la competencia biológica. En el estudio de la competencia biológica, la teoría de juegos proporciona un marco conceptual útil para entender el comportamiento.

Muchos de los juegos más conocidos tienen oponentes o competidores que deben hacer una secuencia de movimientos de acuerdo con las reglas del juego. En algunos juegos, los movimientos sucesivos se hacen con una información completa sobre las oportunidades del oponente (como en el ajedrez). En otros

juegos, los movimientos se hacen con información incompleta (como en el bridge). Un jugador puede decidir sus movimientos al azar (por ejemplo, lanzando una moneda) o de manera deliberada considerando las jugadas posibles. El juego puede terminar después de un número finito de movimientos con un ganador y un perdedor. Generalmente hay un premio al ganador del juego, que puede ser en efectivo o la mera satisfacción de haber ganado. (El premio que obtiene una especie en el juego ecológico es la posibilidad de seguir jugando).

Un juego se caracteriza por sus reglas. En algunos casos, el juego puede ser tan complicado que resultará difícil descubrirlas. Considérese el problema de determinar las reglas del ajedrez viéndolo jugar. Después de cuatro o cinco juegos, se habrán descubierto las reglas principales, pero sería necesario observar muchos más partidos para determinar el resto de las reglas. De manera similar, se pueden considerar las complejas interacciones que se dan en un sistema social humano o de un sistema ecológico como un juego entre muchos jugadores y reglas difíciles de entender de manera completa.

Cuando ya se conocen las reglas de un juego, el problema consiste en determinar la manera que tienen los jugadores de escoger sus movimientos y las consecuencias de estos movimientos, en otras palabras, los jugadores deben determinar sus estrategias analizando las reglas del juego. El resultado final de un juego ordinariamente depende de manera crítica de la selección de movimiento de todos los jugadores. En juegos complejos, puede ser imposible analizar todas las posibilidades, en ese caso, los jugadores deben basarse en la experiencia, la intuición, o en simples pruebas u errores para determinar sus movimientos.

En esta sección se estudiará un juego simple entre dos personas llegando a un nivel considerable de detalle. Los conceptos que se presentan y los resultados deducidos para este juego forman un modelo de análisis para juegos más generales. Empezamos con tres ejemplos de juegos entre dos personas.

Ejemplo 2.18

Monedas emparejadas. Dos jugadores, R y C, colocan cada uno una moneda frente a ellos, cubriéndola, y determinan si el lado expuesto de la moneda es cara (H) o cruz (T). Ninguno de los dos jugadores sabe al principio que lado de la moneda escogió el otro jugador. Entonces se descubren las monedas. Si las dos muestran el mismo lado (las dos son caras o cruces), el jugador R paga \$1 al jugador C. Si las monedas muestran lados distintos, el jugador C paga \$1 al jugador R. Este juego se puede describir en términos de una matriz de pagos.

		Jugador C	
		H	T
Jugador R	H	-1	1
	T	1	-1

Si, por ejemplo, el jugador *R* escoge *A* y el jugador *C* escoge *S*, el jugador *R* gana 1 unidad (en este caso $1). Si los dos jugadores escogen *A*, pierde una unidad el jugador *R*.

Ejemplo 2.19

Un Juego de Negocios. Los únicos dos supermercados en Ciudad Central son Grandes Ahorros y Alimentos El Gigante. El mercado al menudeo se surte de estas dos empresas. Debido al incremento en los costos, Grandes Ahorros quiere aumentar sus precios. Pero teme que si lo hace perderá parte de las ventas en favor de Alimentos El Gigante. Por otra parte, si disminuye sus precios, mientras Alimentos El Gigante aumenta los suyos, el aumento resultante de las ventas compensará con creces las utilidades menores por artículo. Cada empresa tiene tres alternativas: aumentar los precios, mantenerlos sin cambios y reducir los precios.

Grandes Ahorros pueden controlar sus propios precios, pero no tiene control en lo que haga Alimentos El Gigante. Para ayudarle en la decisión, contrata a una analista de mercados independiente, que obtiene los daros de la tabla 1.

Tabla 1. Alternativas de mercadeo para supermercados en competencia

Alternativas para alimentos "El Gigante" / Alternativas para grandes ahorros	A Aumentar precios	M Mantener precios	D Disminuir precios
(A) Aumentar precios	2	2	-7
(M) Mantener precios	6	0	-3
(D) Disminuir precios	10	5	3

Los números de la tabla representan aumentos o disminuciones porcentuales. Por ejemplo, si Grandes Ahorros mantiene sus precios Alimentos el Gigante en favor de Alimentos El Gigante. Si Grandes Ahorros disminuye sus precios

y Alimentos el Gigante en favor de Alimentos El Gigante. Si Grandes Ahorros disminuye sus precios y alimentos El Gigante los aumenta, Grandes Ahorros ganará el 10% del mercado.

Estos datos se pueden representar en la siguiente matriz de pagos.

		Alimentos "El Gigante"		
		A	M	D
Grandes Ahorros	A	2	-2	7
	M	6	0	3
	D	10	5	3

Ejemplo 2.20

Juego de Guerra. Durante la Segunda Guerra Mundial, ocurrió una batalla crítica, la Batalla del Mar de Bismarck, para controlar Nueva Guinea. El jefe de los aliados, el general Kenney, tenía reportes de la inteligencia que indicaban que los japoneses harían movimientos de tropa y convoyes del puerto de Rabaul, en la punta oriental de la isla de Nueva Bretaña, a Lae, que está justo al este de Nueva Bretaña en Nueva Guinea. El jefe de los japoneses tenía dos alternativas: tomar una ruta pasando por el norte de Nueva España, o bien otra por el sur de Nueva Bretaña. En la ruta por el norte, era casi seguro que la visibilidad sería muy mala, mientras que en la ruta por el sur era probable que el clima estuviera despejado. El viaje les tomaría tres días por cualquiera de las rutas.

El general Kenney tenía la opción de concentrar la mayor parte de sus aviones de reconocimiento en una ruta o en la otra. Una vez localizado, el convoy podría ser bombardeado hasta su llegada a Lae. En días de bombardeo, el personal de Kenney estimaba para las distintas opciones los resultados que se dan en la Tabla 2.

Tabla 2 Alternativas para los japoneses y los aliados (número de días bombardeo)

Opciones para los japoneses / Opciones para los aliados	Ruta Norte	Ruta Sur
Ruta Norte	2	2
Ruta Sur	1	3

Una vez más, estas opciones se pueden representar por medio de una matriz de pagos.

		Opciones para los japoneses	
		N	S
Opciones para los aliados	N	2	2
	S	1	3

¿Qué rutas se debieron escoger? Si el convoy japonés tomara la ruta norte, se expondría a 1 o 2 días de bombardeo. Si tomara la ruta sur, se tendría que enfrentar con 2 o 3 días de bombardeo. Parece mejor tomar la ruta norte. Desde el punto de vista del general Kenney, si concentra sus fuerzas en el norte, garantizaría al menos 2 días de bombardeo; en el sur solo podría garantizar 1 día de bombardeo.

Resulta que los dos comandantes escogieron la ruta norte, como se verá próximamente, estas elecciones son consistentes con las previstas por la teoría de juegos.

Estos ejemplos conducen a la siguiente definición.

Definición 2.5.5 **Juego de Matriz**

Sea $A= a_{ij}$ una matriz $m{\times}n$). Considere un juego determinado por A entre dos competidores R y C (renglones y columnas) de acuerdo a las siguientes reglas.

1) En cada movimiento del juego, R escoge uno de los m renglones de A, y C escoge una de las n columnas de A. Estas selecciones se hacen simultáneamente, y ninguno de los jugadores sabe de antemano la elección (o movimiento) del otro competidor.
2) Si R escoge el renglón i y C escoge la columna j, C debe pagar a R la cantidad a_{ij}. Si a_{ij} es negativo quiere decir que C recibe una cantidad $-a_{ij}$ de R.

Este es el juego de matriz $m{\times}n$ determinado por la matriz $m{\times}n$ denotada por $A=(a_{ij})$. El juego de matriz puede terminar después de un movimiento o puede continuar durante cualquier número de movimientos. La matriz del juego se llama matriz $A=(a_{ij})$ del juego o matriz de pagos.

Ejemplo 2.21

Describir los juegos de matriz correspondientes a cada una de las siguientes matrices de pago.

(a) $A=\begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}$

(b) $B=\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 \\ -1 & 0 & 3 & 1 \end{pmatrix}$

a) En este juego de matriz 2×2, *R* y *C* tienen dos opciones cada uno. Si *R* escoge el primer renglón, *R* gana 1 si *C* escoge la primera columna, o 2 unidades si *X* escoge la segunda columna. En el caso de que *R* escoja el segundo renglón, pierde 2 unidades si *C* escoge la primera columna y gana 3 unidades si *C* escoge la segunda columna. Si *C* juega racionalmente, escogerá la primera columna. En este caso, *R* debería escoger el primer renglón. Con estas opciones, *R* tiene la garantía de ganar al menos una unidad y *C* tiene la garantía de no perder más de 1 unidad.
b) En este juego de matriz 3×4, *R* tiene tres opciones y *C* tiene cuatro opciones. Analizando las opciones posibles, es claro que la mejor opción de *C* consiste en escoger la segunda columna. Con esta opción, C tiene la garantía de no perder. La mejor opción de R está en la elección del primer renglón. con estas selecciones de entre las opciones posibles. No se da ningún pago entre los jugadores.

El juego general de matriz $m{\times}n$ es un ejemplo de suma cero entre dos personas, ya que son dos los que compiten y la suma de sus ganancias es cero. Las ganancias de un competidor equivalen a las pérdidas del otro.

Los dos jugadores de un juego de la matriz $A=(a_{ij})$ *dem* $\times$ n de deben analizar sus posibles movimientos y decidir en qué renglones o columnas jugar en movimientos sucesivos. Una estrategia pura para *R*(o *C*) equivale a la decisión de jugar en el mismo renglón (o columna) en cada movimiento del juego. Se dice que el jugador *R*(o *C*) está usando una estrategia mixta si escoge más de un renglón (o columna) en movimientos distintos del juego. Si ambos jugadores usan estrategias puras. El resultado de cada movimiento es exactamente el mismo y el juego es completamente predecible. Por ejemplo, si *R* siempre escoge el renglón *i* y *C* siempre escoge la columna *j,* en cada juego *R* recibe a_{ij} unidades de *C*. Cuando se usan estrategias mixtas por alguno de los jugadores o por ambos, el juego es más complicado. Por ejemplo, si *R* decide jugar una estrategia mixta, hará su selección aleatoriamente de entre los renglones para aumentar sus ganancias.

UN POCO DE TEORÍA DE PROBABILIDADES

Suponga que se lleva a cabo un experimento. A cada resultado posible *E* del experimento, le asignamos un número entre 0 y 1. Este número se llama la probabilidad de que se dé el resultado *E* y se denota por *P*(*E*). Debe enfatizarse que $0 \leq P(E) \leq 1$ Por ejemplo, si se lanza al aire una moneda no cargada, entonces *P*(cara) = *P*(cruz) = ½. Como otro ejemplo, una baraja tiene 52 cartas. De éstas, 13 son corazones. Así que, si se escoge al azar una carta de la baraja, se tiene que:

$$P(\text{corazón}) = \frac{13}{52} = \frac{1}{4}$$

En un experimento, la suma de las probabilidades de todos los posibles resultados es 1. Un vector de probabilidades es un vector con componentes no negativas y cuya suma de las componentes es uno. Los siguientes son vectores de probabilidad: (½, ½), (¼, ½ , 0, ¼) y (0.11, 0.23, 0.17, 0.08, 0.32, 0.09).

Una **variable aleatoria** es una función que asigna un número a cada posible resultado de un experimento. El **valor esperado**, **esperanza o la media**, de un valor aleatorio es un promedio pesado de los valores que se puede tomar la variable aleatoria. Para calcular el valor esperado, se multiplica cada valor de la variable aleatoria por la probabilidad de obtener ese valor. Por ejemplo, supóngase que los valores posibles de la variable aleatoria son 7, 5, –3, y 10, con probabilidades $P(7)=1/8$, $P(5)=5/16$, $P(-3)=1/2$ y $P(10)=1/16$. (Observe que estas probabilidades suman 1.)

Se tiene:

$$\begin{aligned}\text{Valor esperado} &= 7P(7) + 5P(5) + (-3)P(-3) + 10P(10)\\ &= 7(2/16) + 5(5/16) - 3(8/16) + 10(1/16)\\ &= (7+25-24+10)/16 = 25/16\end{aligned}$$

Regresando a nuestra discusión de la estrategia.

¿Cuándo se usará una estrategia pura y cuándo una mixta? Para dar respuesta a estas preguntas necesitamos una mayor precisión sobre lo que entendemos por estrategia.

Definición 2.5.6 **Estrategia**
Una estrategia para R en el juego de la matriz $A=(a_{ij})$ de $m \times n$ es un vector de probabilidad $p=(p_1\ p_2\ \cdots\ p_m)$ donde p_i es la probabilidad de una R juegue escogiendo el renglón i para i= 1,2, ⋯, m. Una estrategia para C un vector de probabilidad de n componentes.

$$q=\begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix}$$

Donde q_j es la probabilidad de que X juegue escogiendo la columna $j = 1, 2, \cdots, n$.

Los jugadores R y C deben escoger sus estrategias $\boldsymbol{p}$ y $\boldsymbol{q}$. En otras palabras, deben escoger las probabilidades p_i y q_j que determinen la frecuencia con la que escogerán los distintos renglones y columnas. Por ejemplo, si R y C escogen el primer renglón y la primera columna de A en todos sus movimientos, están jugando las estrategias puras $\boldsymbol{p}=(1\ 0\ \cdots\ 0)$ y:

$$q=\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Si R y C juegan en todos los renglones y columnas con probabilidades iguales, están jugando las estrategias mixtas $\boldsymbol{p}=(1m\ 1/m\ \cdots\ 1/m)$ y:

$$q=\begin{pmatrix} \frac{1}{n} \\ \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix}$$

Cada vector de probabilidad de m componentes es una posible estrategia para R, y cada vector de probabilidad de n componentes es una posible estrategia para C.

Para ver cuándo se podría usar una estrategia pura, considere el ejemplo siguiente:

$$A = \begin{pmatrix} 6 & 2 & 3 \\ 6 & 5 & 4 \\ 7 & -1 & 2 \\ 2 & 6 & 1 \end{pmatrix}$$

¿Qué estrategia deben adoptar *R* y *C*?

Solución

R jugará de manera que la mínima cantidad que pueda ganar sea tan grande como sea posible (vuelva a leer esto último). Si *R* juega en el renglón 1, ganará al menos 2 unidades, independientemente de qué columna escoja *C*. Si *R* juega en el renglón 2, ganará al menos 4 unidades. De igual forma, si *R* juega en el renglón 3 o en el renglón 4, ganará al menos –1 o 1 unidad, respectivamente.

Así que su mayor ganancia mínima será de 4 unidades.

Pero ¿cómo debe jugar *C*? *C* quiere minimizar su pérdida máxima. Si *C* juega en la columna 1, puede perder hasta 7 unidades; *C* puede perder en la columna 2 a lo más 6 unidades y puede perder en la columna 3 hasta 4 unidades.

Escribimos estos números a continuación:

$$\begin{array}{cccccc} & & C & & & \\ & 6 & 2 & 3 & \leftarrow & 2 \\ R & 6 & 5 & 4 & \leftarrow & 4 \\ & 7 & -1 & 2 & \leftarrow & -1 \\ & 2 & 6 & 1 & \leftarrow & 1 \\ & \uparrow & \uparrow & \uparrow & & \\ & 7 & 6 & 4 & & \end{array}$$

Mínimo del renglón

Máximo de la columna

El número 4 en la posición 2, 3 es un mínimo en su renglón y máximo en su columna. Un número con esas propiedades y máximo en su columna. Un número con esas propiedades se llama un **punto silla** para la matriz de pagos *A*. Cuando un número a_{ij} es un punto silla, las estrategias óptimas para *R* y *C* son,

para *R*, jugar en el renglón *i* y para *C*, jugar en la columna *j*. Así, en el ejemplo, *R* debe adoptar la estrategia pura de jugar en el segundo renglón: $p=(0\ 1\ 0\ 0)$, y *C* debe adoptar la estrategia pura de jugar en la tercera columna

$$q=\begin{pmatrix}0\\0\\1\end{pmatrix}$$

Antes de abandonar este ejemplo, hacemos una observación que puede simplificar los cálculos. Cada número en el primer renglón de *A* es menor o igual que la componente correspondiente en el segundo renglón de *A*. Es decir, $6\leq 6$, $2<5$ y $3<4$. Por lo tanto, *R* nunca tendrá que escoger el primer renglón porque para él la elección del segundo renglón siempre será una opción al menos tan buena como la elección del segundo renglón siempre será una opción al menos tan buena como la elección del primer renglón. El primer renglón constituiría lo que se llama un renglón recesivo.

Similarmente, cada número de la primera columna de es mayor que el número correspondiente en la tercera columna de *A*. Es decir, $6>3$, $6>4$, $7>2$ y $2>1$. Así que *C* nunca escogerá la primera columna, porque si lo hiciera, perdería ciertamente más que si escogiera la tercera columna (recuérdese que la ganancia de *R* es la pérdida de *C*).

La columna 1 es una columna recesiva. Se pueden eliminar los renglones y las columnas recesivos en el análisis subsecuente. Al hacerlo, se obtiene la nueva matriz de pagos A'

$$A'=\begin{pmatrix}5 & 4\\-1 & 2\\6 & 1\end{pmatrix}$$

Igual que antes, 4 es un mínimo en su reglón y un máximo en su columna y por lo tanto un punto silla de A' El segundo renglón es recesivo, por lo que se puede reducir aún más la matriz para obtener

$$A''=\begin{pmatrix}5 & 4\\6 & 1\end{pmatrix}$$

Entonces $A'' = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$ ya que la primera columna de A'' es recesiva. Continuamente de esta manera, se ve que el segundo renglón de A''' es recesivo, por lo que $A^{iv} = (4)$.

Bosquejamos ahora una estrategia general para jugar un juego de matriz en los casos en que se presenta un punto silla.

DETERMINACIÓN DE LAS ESTRATEGIAS PURAS PARA UN JUEGO DE MATRIZ

PASO 1 Eliminar todos los renglones recesivos y todas las columnas recesivas.

PASO 2 Encontrar el número mínimo en cada renglón. Este se llama mínimo del renglón.

PASO 3 Encontrar el número máximo en cada columna. Este se llama máximo de la columna.

PASO 4 Buscar un punto silla. Este es un número que es a la vez un mínimo de renglón y un máximo de columna. Si aij es un punto silla, R deberá jugar en el renglón *i* y *C* deberá jugar en la columna j. En este caso se dice que el juego de matriz está determinado estrictamente.

PASO 5 Si no hay ningún punto silla, *R* o *C* (o ambos) debe usar una estrategia mixta. El juego en ese caso no está determinado estrictamente.

Ejemplo 2.22

Determine si el juego definido por la matriz de pagos dada en los siguientes incisos es estrictamente determinado. En caso afirmativo, determinar las estrategias óptimas para *R* y *C*.

$$\text{a)} \begin{pmatrix} 3 & 0 & 5 \\ 2 & 1 & 3 \\ 2 & -1 & -2 \end{pmatrix} \quad \text{b)} \begin{pmatrix} 1 & 6 & 2 \\ 4 & 2 & 3 \\ 5 & 1 & 6 \end{pmatrix} \quad \text{c)} \begin{pmatrix} 1 & 0 & 5 \\ 1 & -5 & 0 \end{pmatrix}$$

Solución

a) Ya que cada número en el renglón 3 es menor o igual que el número correspondiente en el renglón 2, el renglón 3 es recesivo. De manera semejante, cada número en la columna 1 es mayor que el número correspondiente en la columna 2, por lo que la columna 1 es recesiva. Eliminando el renglón 3 y la columna 1, se obtiene:

$$\begin{pmatrix} 0 & 5 \\ 1 & 3 \end{pmatrix} \begin{matrix} \Leftarrow & 0 \\ \Leftarrow & 1 \end{matrix} \quad \text{Mínimo del renglón}$$

$$\begin{matrix} & \Uparrow & \Uparrow \\ \text{Máximo de la columna} & 1 & 5 \end{matrix}$$

Puede verse que 1 es a la vez mínimo en su renglón y máximo en su columna, por lo que 1 es un punto silla y el juego es estrictamente determinado. Como 1 es la componente 2, 2 de la matriz de pagos, las estrategias óptimas son $\boldsymbol{p}= (0\ 1\ 0)$ y

$$q=\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Es decir, *R* escoge el renglón 2 y *C* escoge la columna 2. *

b) No existen renglones ni columnas recesivos. Volviendo a escribir la matriz de pagos se tiene:

$$\begin{pmatrix} 1 & 6 & 2 \\ 4 & 2 & 3 \\ 5 & 1 & 6 \end{pmatrix} \begin{matrix} 1 \\ 2 \\ 1 \end{matrix} \Leftarrow \text{Mínimo del renglón}$$

$$\text{Máximo de la columna} \Rightarrow \quad 5 \quad 6 \quad 6$$

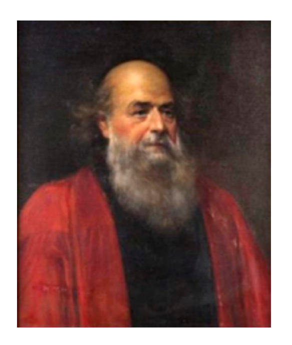
En este caso no hay ningún punto silla porque no existe un número que sea a la vez un mínimo en su renglón y un máximo en su columna. El juego no está estrictamente determinado y se requiere usar una estrategia mixta.

c) El renglón 2 es recesivo, así como la columna 1. La matriz de juego se reduce a $(0\ \ 5)$ en donde 0 es un punto silla. Obsérvese que la matriz de pagos tiene dos ceros, pero tiene dos ceros, pero solo uno de ellos, el que está en la posición 1, 2, es un punto silla. El juego está estrictamente determinado y las estrategias óptimas son $\boldsymbol{p}(1\ \ 0)$ y:

$$q=\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Así que *R* juega en el renglón 1 y *C* juega en la columna 2.

2.6. NOTAS HISTÓRICAS



El término **matriz** se mencionó por primera vez en la literatura matemática en un artículo de 1850 de James Joseph Sylvester (1814-1897). El significado usual no técnico de ese término es *"lugar donde algo se crea, produce o desarrolla"*. Para Sylvester, entonces, una matriz (definida como "un ordenamiento oblongo de términos") era una entidad a partir de la cual uno podía formar varias porciones cuadradas para producir determinantes. Estas últimas cantidades, formadas a partir de matrices cuadradas, eran bastante bien conocidas en esa época.



La **multiplicación de matrices** se originó en la composición de sustituciones lineales. Aunque hay indicios de dichas ideas en el trabajo de Euler y Lagrange, fue Karl Friedrich Gauss (1777-1855) quien las estudió de modo exhaustivo en su trabajo de 1801, *DisquisitionesArithmeticae*, en conexión con formas cuadráticas (funciones con dos variables de la forma $Ax2 + 2Bxy + Cy2$). En particular, una sustitución lineal de la forma:

$$x = ax' + by'\ y = cx' + dy' \qquad (1)$$

Convierte una de dichas formas F en x e y; en otra forma, F' en x', y'. Si una segunda sustitución:

$$x' = ex + fy' = gx'' + hy'' \qquad (2)$$

Transforma F' en una forma F'' en x'', y'' entonces la composición de las sustituciones, hallada al reemplazar x', y' en (1) por sus valores en (2), da una sustitución que transforma F en F'':

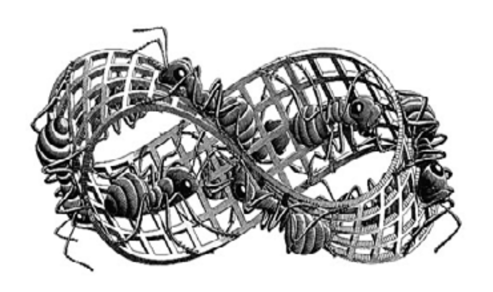
$$x = (ae+bg)x + (af+bh)y = (ce+dg)x'' + (cf+dh)y'' \qquad (3)$$

La matriz de coeficientes de la sustitución (3) es el producto de las matrices de coeficientes de las sustituciones (1) y (2). Gauss realizó una composición

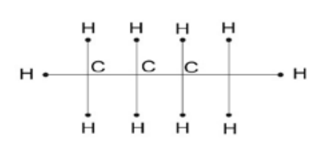
análoga de sustituciones para formas con 3 variables, que da la multiplicación de matrices de 3×3.

A Gauss se le suele llamar “príncipe de las matemáticas”, pues durante una larga carrera científica hizo importantes contribuciones a campos tan variados como teoría de los números, álgebra, geometría, análisis complejo, astronomía, geodesia y mecánica.

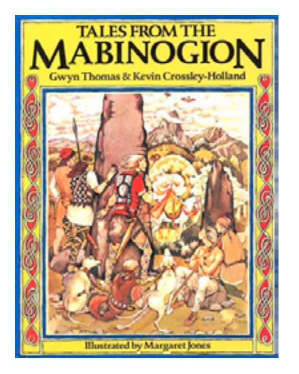
TEORÍA DE GRÁFICAS



La rama de las matemáticas que estudia estos objetos se conoce como teoría de gráficas, y surgió en el siglo XVIII, época en que los matemáticos veían en las gráficas un simple divertimento. Es por eso que no resulta sorprendente que gran parte de los resultados obtenidos inicialmente en esta área hayan surgido a partir de acertijos y pasatiempos. Fue hasta la última mitad del siglo pasado cuando el interés por la teoría de gráficas se incrementó notablemente. Desde luego, las razones de tal fenómeno son muchas, aunque otra vez su versatilidad parece ser determinante: se ha encontrado una enorme diversidad de aplicaciones de la teoría de gráficas en muchas áreas del conocimiento. Así, la vemos haciendo aportaciones en biología, ingeniería civil, arquitectura, genética, economía, antropología, lingüística, química, física, economía y por supuesto en otras áreas de las matemáticas.



En 1857 **Cayley** descubrió una importante clase de gráficas, llamadas árboles en la química orgánica y las usó para representar los isómeros de los hidrocarbonos saturados. De este modo Cayley descubrió cuántos isómeros (es decir cuántos árboles) existen para un número dado de átomos de carbono.



Historia de la Teoría de Juegos. La noción más antigua de un juego se encuentra en el Mabinogion, una colección de cuentos populares galeses (siglos XI-XIII). Hay un relato en el que dos reyes que están en guerra

juegan al ajedrez, mientras que sus ejércitos batallan en las proximidades. Cada vez que un rey se come una pieza llega un mensajero para informar al otro que ha perdido un hombre importante o una división. Al final un rey da jaque al otro. http://herzog.economia.unam.mx/profesores/blopez/juegos-Introducci%C3%B3n.pdf



John Von Neumann (1903-1957) Fue un matemático húngaro-estadounidense que realizó contribuciones fundamentales en física cuántica, análisis funcional, teoría de conjuntos, ciencias de la computación, economía, análisis numérico, cibernética, hidrodinámica, estadística y muchos otros campos. Desde la década de 1920 estuvo trabajando en la estructura matemática del póker y otros juegos, pero enseguida vio que sus teoremas podían ser aplicados a economía, política, relaciones internacionales, etc. No fue hasta 1944, cuando Von Neumann y Morgensten, publicaron su libro Teoría de Juegos y Comportamiento Económico, que incide en el desarrollo de la programación lineal y la teoría de la decisión estadística de Wald. John Von Neumann demostró matemáticamente que siempre hay un curso racional de acción para juegos de dos jugadores, con intereses completamente opuestos (uno gana y el otro pierde). Esta prueba es conocida como el Teorema Minimax. http://herzog.economia.unam.mx/profesores/blopez/juegos-Introducci%C3%B3n.pdf



Émile Borel (1871-1956) En 1921 el matemático francés, EB, hizo públicos varios artículos sobre la théorie du jeu ("Game theory and left symetric core integral equations"). Borel usó el póquer como ejemplo, y analizó el problema del faroleo. Borel reparó en las posibles aplicaciones económicas y militares de la teoría de juegos.

Planteó las cuestiones esenciales de la teoría de juegos: ¿para qué juegos existe la mejor estrategia, y de qué manera puede uno buscar esa estrategia? Como matemático, era conocido por su trabajo fundación en las áreas de la teoría de la medida y la probabilidad. http://herzog.economia.unam.mx/profesores/blopez/juegos-Introducci%C3%B3n.pdf



John Forbes Nash (1928-2015), Economista y matemático estadounidense. Extraordinariamente dotado para el análisis matemático, Nash desarrolló investigaciones en torno a la teoría de juegos, que le valieron el Premio Nobel de Economía en 1994, junto a John Harsanyi y Reinhard Selten.

Ingresó en el Carnegie Institute of Technology, en la actualidad Universidad Carnegie-Mellon de Pittsburgh, con la intención de estudiar ingeniería química; pero tras cursar algunas asignaturas de matemáticas, aceptó la sugerencia de sus profesores de orientar su carrera hacia esta materia. En 1948 obtuvo el grado de licenciado en matemáticas y tras recibir varias ofertas para realizar el doctorado, se decidió por la Universidad de Princeton.

A lo largo de sus estudios doctorales mostró interés por diversos campos de estudio, como la topología, el álgebra geométrica o la teoría de juegos. En 1949 y como parte de sus investigaciones publicó en la revista Annals of Mathematics un artículo titulado “Non-cooperative Games”, en el que se recogían las ideas principales de su tesis, que presentó el siguiente año en Princeton. En dicho artículo se exponían los puntos básicos sobre las estrategias y las posibilidades de predicción del comportamiento que se da en juegos no cooperativos con información incompleta.

Una vez finalizada su tesis, trabajó durante unos meses para la Corporación RAND, que estaba muy interesada en sus conocimientos de la teoría de juegos para aplicarlos a la estrategia militar y diplomática. Volvió a la Universidad de Princeton poco después, lo que no resultó impedimento para que colaborara de forma esporádica con la Corporación RAND. En 1952 se incorporó al cuerpo docente del prestigioso Massachusetts Institute of Technology (MIT), donde realizó una importante labor de investigación sobre variables algebraicas reales múltiples.

2.7. EJERCICIOS PROPUESTOS

1. Sean A y B matrices de orden 4×4 cuyos valores desconocemos, pero se nos dice que $DA=DB$ para toda matriz diagonal ¿Es cierto que $A=B$? Justifique su respuesta.

2. a) Demostrar que, si A es una matríz y x es un vector renglón, entonces xA, si está definido es de nuevo un vector renglón.
   b) Demostrar que, si A es una matríz e y es un vector columna, entonces Ay, si está definido es un vector columna.

3. Calcule los productos de las siguientes matrices:

a) $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \Gamma & \delta \end{bmatrix}$

b) $\begin{bmatrix} 1 & -3 & 2 \\ 3 & -4 & 1 \\ 2 & -5 & 3 \end{bmatrix} \begin{bmatrix} 2 & 5 & 6 \\ 1 & 2 & 5 \\ 1 & 3 & 2 \end{bmatrix}$

c) $\begin{bmatrix} \cos a & -\text{sen}\, a \\ \text{sen}\, a & \cos a \end{bmatrix}^n$

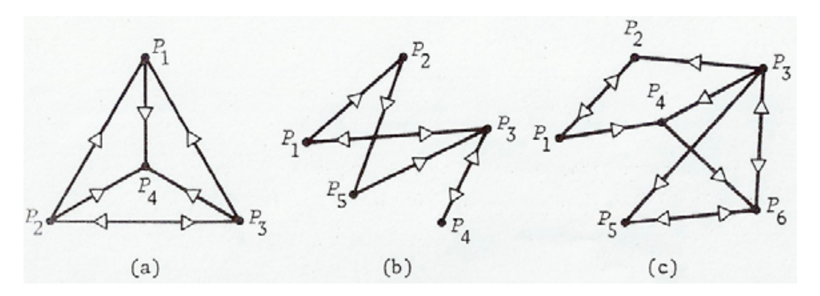
d) $\begin{bmatrix} a_i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n \end{bmatrix}^k$

e) $A=\begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}^3$

Donde A es de orden n.

4. ¿Cómo se altera el producto AB de las matrices A y B si
   a) Se intercambia la i-ésima y j-ésima filas de la matríz A?
   b) A la i-ésima fila de la matríz se le añade la j-ésima fila multiplicada por el escalar C?

5. Construir la matriz de los vértices de cada una de las gráficas dirigidas que aparecen en la figura siguiente:



6. Trazar un diagrama de una gráfica dirigida que corresponda a cada una de las siguientes matrices de los vértices.

(a) $\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$

7. Sea M la siguiente matriz de los vértices de una gráfica dirigida:

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

a) Trazar un diagrama de la gráfica.
b) Aplicar el teorema 2.1 para obtener el número de ligas de 1, 2 y 3 pasos que unen al vértice P_1 con el vértice P_2. Verificar la respuesta listando todas las ligas, como se hizo en el ejemplo 2.3.
c) Repetir (b) para ligas de 1, 2 y 3 pasos que conectan a P_1 con P_4.

Capítulo 3
SOLUCIÓN DE SISTEMAS LINEALES

3.1. INTRODUCCIÓN

A menudo se compara la tarea de los matemáticos con la de los criminalistas que, forzosamente, tenían que hacer prender una incógnita. Hay, sin embargo, un aspecto en el que esta comparación nos deja abandonados en el problema: cualquier tribunal reputará de mucho más difícil la captura de un malhechor solitario, que solo se confía a sí mismo prescindiendo de cómplices peligrosos; una gran banda de ladrones se traiciona a sí misma mucho antes. En las matemáticas y en particular en el álgebra lineal, por el contrario, cada nueva incógnita nos hace la vida considerablemente más difícil.

Imaginemos para empezar, con un ejemplo numérico sencillo:

$$x + y = 62$$

x e y son dos números desconocidos, que precisamente tratamos de encontrar. Pero, evidentemente, esta exigencia está completamente indeterminada, puesto que hay toda una serie de números que satisfacen a la ecuación, como, por ejemplo: x=1, y=61; x=6, y=56; x=70, y=–8, por no hablar de la avalancha de fracciones. Hay pues una infinidad de pares de números, x, y que representan soluciones a la ecuación; la cosa tiene así poca apariencia de problema matemático.

El caso recuerda aquella otra pregunta: Entre mis amistades hay un matrimonio que entre los dos suman 62 años. ¿Qué edad tiene cada uno de ellos? La cuestión cambia enseguida cuando se agrega una segunda condición, por ejemplo, que la diferencia de sus edades es de dos años. Con ello el problema queda determinado y se tiene un sistema de ecuaciones lineales. Este problema o ejemplo tan simple da lugar a un sistema, mismo que podemos encontrar en

problemas con más incógnitas y más ecuaciones referidas por ejemplo a restricciones industriales.

En este capítulo analizaremos más de cerca lo referente a sistemas de ecuaciones lineales y sus soluciones.

3.2. EL MÉTODO DE REDUCCIÓN DE GAUSS

Recordemos nuevamente que se busca resolver sistemas de ecuaciones lineales cuadrados, es decir: con el mismo número de ecuaciones que de incógnitas, en este sentido el método de reducción de Gauss ha mostrado ser altamente eficiente y a continuación se presenta.

Sea el sistema:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{3.1}$$

Este sistema está completamente determinado por su matriz de coeficientes $A=(a_{ij})$ y el vector columna b con i-ésimo elemento b_i. La matriz aumentada se denota (A/b) y se escribe:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{2n} & b_2 \\ a_{n1} & a_{n2} & a_{nn} & b_n \end{array}\right] \tag{3.2}$$

El vector columna

$$x=\begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix}$$

Cuyas componentes satisfacen el sistema (3.1) es una solución de dicho sistema.

Es fácil ver que las soluciones del sistema (3.1) son las mismas que las de cualquier sistema obtenido de (3.1) por los siguientes procedimientos.

1. Intercambio de dos ecuaciones.
2. Multiplicación de una ecuación del sistema (3.1) por una constante ≠0.
3. Sustitución de una ecuación por la suma de sí misma con un múltiplo de una ecuación diferente del sistema.

Aplicados al sistema (3.1), estos procedimientos corresponden a las **operaciones elementales en las filas** que se aplican a toda la matriz aumentada (3.2).

Para resolver el sistema (3.1) mediante el método de Gauss, tratamos de usar operaciones elementales en las filas para reducir la matriz (3.2) a una matriz de la forma:

$$\left[\begin{array}{ccc|c} u_{11} & u_{12} & u_{1n} & c_1 \\ & u_{22} & u_{2n} & c_2 \\ & \cdots & \cdots & \cdots \\ & & u_{nn} & c_n \end{array}\right]$$

Donde la parte cuadrada U a la izquierda de la partición tiene entradas igual a cero debajo de la diagonal principal. La matriz de coeficientes A original del sistema (3.1) se transforma en una **matriz triangular superior *U*** con ceros debajo de la diagonal principal. El sistema se convierte en:

$$Ux = c$$

Si una matriz B se puede obtener de una matriz A a través de operaciones elementales en las filas, entonces B es equivalente por filas a A. Por lo cual las matrices A y U descritas antes son equivalentes por filas.

Ejemplo 3.1

Si se tiene la matriz triangular superior aumentada:

$$\left[\begin{array}{ccc|c} -5 & -1 & 3 & 3 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 2 & -4 \end{array}\right]$$

Encuentra la solución del sistema.

Solución

Las ecuaciones correspondientes serían:

$$-5_1 - x_2 + 3x_3 = 3$$
$$3x_2 + 5x_3 = 8$$
$$2x_3 = -4$$

De la última ecuación se tiene que $x_3 = -2$; sustituyendo en la segunda ecuación se tiene:

$$3x_2 + 5(-2) = 8$$

Y de aquí $\quad x_2 = 6$

Sustituyendo x_2 y x_3 en la primera ecuación se obtiene:

$$-5x_1 - 6 + 3(-2) = 3$$

Y de aquí: $\quad x_1 = -3$

Así, la solución es $x^T = [-3, 6, -2]$. Este procedimiento recibe el nombre de sustitución regresiva.

Ejemplo 3.2

Resuelva el sistema lineal a través del método de Gauss con sustitución regresiva:

$$x_2 - 3x_3 = -5$$
$$2x_1 + 3x_2 - x_3 = 7$$
$$4x_1 + 5x_2 - 2x_3 = 10$$

Solución

Se reduce la matriz aumentada por medio de operaciones elementales en las filas. Los pivotes se encierran en un círculo.

Se intercambian las filas 1 y 2:

$$\left[\begin{array}{ccc|c} 0 & 1 & -3 & -5 \\ 2 & 3 & -1 & 7 \\ 4 & 5 & -2 & 10 \end{array}\right] \approx \left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 4 & 5 & -2 & 10 \end{array}\right]$$

CONTENIDO

Se suma la fila 1 multiplicada por –2 a la fila 3:

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 4 & 5 & -2 & 10 \end{array}\right] \approx \left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & 0 & -4 \end{array}\right]$$

Se suma la fila 2 a la 3:

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & 0 & -4 \end{array}\right] \approx \left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & -3 & -9 \end{array}\right]$$

De la última matriz obtenemos, por sustitución regresiva:

$x_3 = 3$ $x_2=4$ $x_1= -1$

Reducción de Gauss a la forma triangular superior:
Caso de solución única

PASO 1 Si la entrada superior de la columna 1 es cero, entonces efectuar una operación de intercambio de filas para obtener un elemento distinto de cero en la parte superior de la columna. Siempre es posible esto en el caso de solución única. Llamamos pivote al elemento elegido distinto de cero.

PASO 2 Efectuar operaciones elementales en las filas de tal forma que las filas inferiores resultantes tengan cero como primer entrada.

PASO 3 Después del paso 2, la matriz tiene la forma:

$$\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} & c_1 \\ 0 & x & \cdots & x & x \\ 0 & \cdots & \cdots & \cdots & \cdots \\ 0 & x & \cdots & x & x \end{bmatrix} \qquad (3.3)$$

Y la primera columna está en la forma que queremos. Se elimina (mentalmente) la fila superior y la primera columna de la matriz (3.3), dejando la parte enmarcada de ésta. Se vuelve al paso 1 con esta matriz más pequeña y se repite el

procedimiento para componer la siguiente columna. Se continúa hasta obtener la forma triangular superior.

3.3. EL MÉTODO DE GAUSS-JORDAN

Una matriz cuadrada está en la forma diagonal si todas las entradas fuera de la diagonal principal son cero. El método de Gauss-Jordan es un avance del método de Gauss y usa operaciones elementales en las filas para reducir la parte izquierda de una matriz aumentada a la forma diagonal con todos los pivotes iguales a uno, es decir, a la matriz identidad o idéntica**I**.

Suponga, por ejemplo, que una matriz aumentada se reduce a la forma:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 9 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

Se ve de inmediato que la solución al sistema original está dada por:

$x_1 = 9$ $\qquad$ $x_2 = 7$ $\qquad$ $x_3 = -1$

Ejemplo 3.3

Encuentre, mediante la reducción de Gauss-Jordan, la solución del sistema:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 7 \\ 2x_1 - 5x_2 + 2x_3 &= 6 \\ 3x_1 + 2x_2 - x_3 &= 1 \end{aligned}$$

Solución

1. Se suma la fila 1, multiplicada por –2, a la fila 2 y se suma la fila 1, multiplicada por –3, a la fila 3:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 2 & -5 & 2 & 6 \\ 3 & 2 & -1 & 1 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & -1 & 0 & -8 \\ 0 & 8 & -4 & -20 \end{array}\right]$$

2. Se multiplica la fila 2 por −1:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & -1 & -0 & 8 \\ 0 & 8 & -4 & -20 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & 1 & 0 & 8 \\ 0 & 8 & -4 & -20 \end{array}\right]$$

3. Se suma la fila 2, multiplicada por 2 a la fila 1 y se suma la fila 2, multiplicada por −8 a la fila 3:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & 1 & 0 & 8 \\ 0 & 8 & -4 & -20 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & -4 & -84 \end{array}\right]$$

4. Se multiplica la fila 3 por −¼

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & -4 & -84 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{array}\right]$$

5. Se suma la fila 3, multiplicada por −1, a la fila 1:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{array}\right]$$

De aquí se observa que la solución es:

$x_1=2$ $x_2=8$ $x_3=21$.

3.4. SISTEMAS CON LA MISMA MATRIZ DE COEFICIENTES

Dados dos sistemas lineales $Ax=b$ con la misma matriz de coeficientes A, pero distintos vectores columna b, suponga que queremos resolver los dos sistemas simultáneamente:

$Ax=b$ $Ay = b'$

En lugar de resolver uno después de otro, es más eficiente tomar una sola matriz aumentada con la matriz de coeficientes A a la izquierda de la partición y aumentarla con dos vectores columna b y b', a la derecha de la partición. Entonces se reduce esta matriz partida de modo que la parte izquierda tenga forma triangular superior o diagonal, resolviendo así ambos sistemas simultáneamente.

Ejemplo 3.4

Resuelva los sistemas lineales utilizando el método de Gauss-Jordan.

$$2x_1 - 4x_2 = -10$$
$$x_1 - 3_{x2} + x_4 = -4$$
$$x_1 - x_3 + 2x_4 = 4$$
$$3x_1 - 4x_2 + 3x_3 - x_4 = -11$$

$$2y_1 - 4y_2 = -8$$
$$y_1 - 3y_2 + y_4 = -2$$
$$y_1 - y_3 + 2y_4 = 9$$
$$3y_1 - 4y_2 + 3y_3 - y_4 = -15$$

Solución

$$\left[\begin{array}{cccc|cc} 2 & -4 & 0 & 0 & -10 & -8 \\ 1 & -3 & 0 & 1 & -4 & -2 \\ 1 & 0 & -1 & 2 & 4 & 9 \\ 3 & -4 & 3 & -1 & -11 & -15 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & -2 & 0 & 0 & -5 & -4 \\ 1 & -3 & 0 & 1 & -4 & -2 \\ 1 & 0 & -1 & 2 & 4 & 9 \\ 3 & -4 & 3 & -1 & 4 & -15 \end{array}\right]$$

$$\left[\begin{array}{cccc|cc} 1 & -2 & 0 & 0 & -5 & -4 \\ 0 & -1 & 0 & 1 & 1 & 2 \\ 0 & 2 & -1 & 2 & 9 & 13 \\ 0 & 2 & 3 & -1 & 4 & -3 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & -2 & 0 & 0 & -5 & -4 \\ 0 & 1 & -1 & -1 & -1 & -2 \\ 0 & 2 & -1 & 2 & 9 & 13 \\ 0 & 2 & 3 & -1 & 4 & -3 \end{array}\right]$$

$$\left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 0 & -1 & 4 & 11 & 17 \\ 0 & 0 & 3 & 1 & 6 & 1 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 0 & 1 & -4 & -11 & -17 \\ 0 & 0 & 3 & 1 & 6 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 0 & 1 & -4 & -11 & -17 \\ 0 & 0 & 0 & 13 & 39 & 52 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 1 & 0 & 1 & -4 & -11 & -17 \\ 0 & 0 & 0 & 1 & 3 & 4 \end{array}\right]$$

$$\left[\begin{array}{cccc|cc} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 3 & 4 \end{array}\right]$$

A partir de esta última matriz vemos que las soluciones son:

$x_1= -1$ $\quad x_2=2$ $\quad x_3=1$ $\quad x_4=3$
$y_1= 0$ $\quad y_2= 2$ $\quad y_3= -1$ $\quad y_4= 4$

De aquí se puede concluir un resultado que se resume en el siguiente teorema:

Teorema 3.1
Sea $Ax=b$ un sistema de n ecuaciones lineales en n incógnitas. El sistema tiene una solución única si y solo si la matriz A de los coeficientes es equivalente en filas a la matriz identidad I de $n{\times}n$.

3.5. MATRICES ELEMENTALES

Una manera de visualizar las operaciones elementales es a través de sus matrices, recordemos que las operaciones fila en una matriz son:

a) Intercambio de filas.
b) Multiplicación de una fila por una constante diferente de cero.
c) Sumar el múltiplo de una fila a otra.

Asimismo, las operaciones columna de una matriz son las mismas que las operaciones fila. La colección de operaciones fila y columna en una matriz se dicen **operaciones matriciales elementales**.

Definición 3.1
Una matriz E se dice elemental si resulta de efectuar una operación elemental en la matriz identidad I de orden $n \times n$.

Las matrices elementales son:

E_{pq} = intercambio de filas (columnas) p y q.
$E_p^{(\alpha)}$ = multiplicar una fila (columna) p por $\alpha \neq 0$.
$E_{pq}^{(\beta)}$ = multiplicar la fila p por β y sumarla a la fila q.

Ejemplo 3.5

Sea

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Considere las siguientes matrices elementales de 3×3:

a) Multiplicar la segunda fila de I_3 por −3:

$$E_2^{(-3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

b) Intercambiar las filas 2 y 3:

$$E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

* **Nota**: $E_{32} = E_{23}$

c) Multiplicar la fila 1 por 4 y sumarla a la fila 3:

$$E_{13}^{(4)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

Observe que para:

$$E_{32}\,A=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{31}&a_{32}&a_{33}\\a_{21}&a_{22}&a_{23}\end{bmatrix}$$

Se intercambian las filas dos y tres de la matriz A.

Si se hace $A\,E_{32}$ se tiene:

$$AE_{32}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}=\begin{bmatrix}a_{11}&a_{13}&a_{12}\\a_{21}&a_{23}&a_{22}\\a_{31}&a_{33}&a_{32}\end{bmatrix}$$

Y se intercambian las columnas dos y tres de la matriz A.

De aquí podemos concluir que: Para efectuar una transformación elemental de fila sobre una matriz A de $n \times n$ se aplica la transformación a la matriz I de $n \times n$ obteniéndose la matriz elemental E correspondiente y a continuación se multiplica por la izquierda A por E; equivalentemente, para efectuar una transformación elemental de columna sobre una matriz A, se multiplica A, por la derecha, por la matriz E correspondiente.

3.6. USO DE MATRICES ELEMENTALES

La reducción de una matriz cuadrada a la forma triangular superior U o diagonal se puede lograr a través de multiplicaciones sucesivas por la izquierda por matrices elementales en las filas, entonces existen matrices elementales E_1, E_2, $\cdots$, E_t tales que:

$$U = (E_t \cdots E_2\, E_1)\, A$$

Ejemplo 3.6

Considere las matrices:

a) $A=\begin{bmatrix} 2 & 4 & 1 \\ 3 & 2 & 2 \\ 1 & 1 & 3 \end{bmatrix}$ $\quad B=\begin{bmatrix} 4 & 4 & 1 \\ 6 & 2 & 2 \\ 2 & 1 & 3 \end{bmatrix}$

b) $A=\begin{bmatrix} 1 & 4 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$ $\quad B=\begin{bmatrix} 1 & 4 & 1 \\ 0 & 1 & 1 \\ 0 & -4 & 1 \end{bmatrix}$

Especifique si existe una matriz elemental E tal que $AE=B$ o $EA=B$.

Solución

a) $AE_1^{(2)}=B$ o $BE_1^{(½)}=A$

b) $E_{13}^{(-1)}A=B$ o $E_{13}^{(1)}B=A$

Ejemplo 3.7

Sea

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = A$$

Encuentre una matriz C tal que CA sea una matriz triangular superior equivalente en filas a A.

Solución

Sean: $E_1=E_{12}$ $\quad E_2=E_{13}^{(-2)}$ $\quad E_3=E_{23}^{(1)}$

Entonces: $C=E_3\,E_2\,E_1$

3.7. MATRICES INVERSAS

Hasta este punto hemos hecho uso de manera implícita de las matrices inversas sin embargo es importante considerar que para que una matriz de $n\times n$ tenga inversa debe ser regular y que no tenga renglones o columnas iguales a cero, más adelante se dan las condiciones de su existencia.

Definición 3.2
Sea A una matriz de $n{\times}n$. Una matriz C de $n{\times}n$ es una inversa de A si $CA{=}AC{=}I$, donde I es la matriz identidad de $n{\times}n$.

Teorema 3.2: Unicidad de la inversa
Sea A una matriz de $n{\times}n$ con inversa C de modo que $CA{=}AC{=}I$. Si D es una matriz de $n{\times}n$ tal que $AD{=}I$, entonces $C{=}D$.

Demostración
Como el producto de matrices es asociativo, se tiene:

$$C(AD) = (CA)D$$

Pero $AD = I$ y $CA = I$; entonces:

$$C(AD) = CI = C \quad \text{y} \quad (CA)D = ID = D$$

De donde: $C = D$.
La inversa de una matriz A se denota por A^{-1}.

Definición 3.3
Una matriz cuadrada que tiene inversa se llama **invertible**. Si tal matriz cuadrada no tiene inversa se llama **singular**.

Ejemplo 3.8
Determine la inversa de la matriz A.

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solución

$$AC = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Haciendo el producto de AC e igualando a I se tiene:

$c_{11}=½ \; c_{12}=c_{13}=0$
$c_{22}=¼ \; c_{21}=c_{23}=0$
$c_{33}=1 \;\; c_{31}=c_{32}=0$

De donde:

$$A^{-1}=C=\begin{bmatrix} ½ & 0 & 0 \\ 0 & ¼ & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ejemplo 3.9

Determine la inversa de la matriz B, donde B_i son matrices cuadradas:

$$B=\begin{bmatrix} B_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & B_3 \end{bmatrix}$$

Solución

Sea C la inversa de B, entonces:

$$BC=\begin{bmatrix} B_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & B_3 \end{bmatrix}\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}=\begin{bmatrix} In_1 & 0 & 0 \\ 0 & In_2 & 0 \\ 0 & 0 & In_3 \end{bmatrix}$$

De aquí se tiene:

$B_1c_{11}=I_{n1}$
$B_2c_{22}=I_{n2}$
$B_3c_{33}=I_{n3}$

Se despejan las c_i's y se obtiene:

$$B^{-1}=C=\begin{bmatrix} B_1^{-1} & 0 & 0 \\ 0 & B_2^{-1} & 0 \\ 0 & 0 & B_3^{-1} \end{bmatrix}$$

Ejemplo 3.10
Sea la matriz A como sigue:

$$A=\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Donde $a\neq 0$, $b\neq 0$ y $c\neq 0$. Calcule la inversa de A.

Solución
Se puede expresar A como producto de matrices A_1, A_2 y A_3 tales que:

$$A=\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}=\underbrace{\begin{bmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{A_1}\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{A_2}\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}}_{A_3}$$

Entonces A^{-1} se puede ver como:

$$\underbrace{\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}}_{A^{-1}}=\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}}_{A_3^{-1}}\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{A_2^{-1}}\underbrace{\begin{bmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{A_1^{-1}}$$

PROPOSICIÓN 3.7.1

Sean A y B matrices de $n\times n$ cuyas inversas son conocidas, entonces:

a) $(A^{-1})^{-1} = A$ (significa que A^{-1} tiene inversa)
b) $(AB)^{-1} = B^{-1}A^{-1}$ (significa que AB tiene inversa)

Demostración
Se demostrará el inciso b); el inciso a) se deja al estudiante como ejercicio.

b) Se da por supuesto que existen matrices A^{-1} y B^{-1} tales que $AA^{-1}=I$ y $BB^{-1}=I$. Utilizando la ley asociativa para producto de matrices, tenemos que:

CONTENIDO

$$(AB)(B^{-1}A^{-1}) = \left[A(BB^{-1})\right] A^{-1} = (AI)A^{-1} = AA^{-1} = I$$

En forma similar:

$$(B^{-1}\, A^{-1})(AB) = \left[B^{-1}\, (A^{-1}A)\right]B = \left[B^{-1}\, (I)\right]B = BB^{-1} = I$$

Por lo tanto: $(AB)^{-1} = B^{-1}A^{-1}$

INVERSAS DE MATRICES ELEMENTALES

Sea E_{ik} una matriz elemental de intercambio de renglones obtenida al intercambiar los renglones *ik*. Recuerde que $E_{ik}A$ afecta intercambiando los renglones *ik* de *A*; donde *A* es cualquier matriz del mismo tamaño que E_{ik}. En particular al tomar $A=E_{ik}$ vemos que $E_{ik}\,E_{ik}$ intercambia las filas *ik* de *E*, y por lo tanto, devuelve a E_{ik} a *I*, así:

$$E_{ik}\,E_{ik} = I$$

Si multiplicamos el renglón *i* por un escalar *r* distinto de cero, entonces:

$$E_i^{(r)}\, E_i^{(1/r)} = I$$

Si se tiene $E_{pq}{}^{(\beta)}$ entonces su inversa está dada por:

$$E_{pq}{}^{(\beta)}E_{pq}{}^{(-\beta)} = I$$

Por lo tanto: ***toda matriz elemental es invertible***.

Ejemplo 3.11

Encuentre las inversas de las siguientes matrices elementales:

$$E_{12}= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_1{}^{(3)} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_{31}{}^{(4)} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solución

$$[E_{12}]^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$[E_1^{(3)}]^{-1} = \begin{bmatrix} ⅓ & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$[E_{31}^{(4)}] = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

CÁLCULO DE INVERSAS

Sea $A=(a_{ij})$ una matriz de $n\times n$. Para hallar A^{-1}, si es que existe, se debe encontrar una matriz $X=(x_{ij})$ de$n\times n$ tal que $AX = I$, esto es, tal que:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{ln} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{ln} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \qquad (3.4)$$

La ecuación matricial (3.4) corresponde a n^2 ecuaciones lineales con las n^2 incógnitas x_{ij}, hay una ecuación lineal para cada una de las n^2 entradas en una matriz de $n\times n$. Por ejemplo, si se igualan las entradas de la posición de la segunda fila y primera columna a cada lado de la ecuación (3.4) se obtiene la ecuación lineal:

$$a_{21}\, x_{11} + a_{22}\, x_{21} + \cdots + a_{2n}\, x_{n1} = 0$$

De estas n^2 ecuaciones lineales, n de ellas comprenden las n incógnitas x_{i1} para $i=1,2,\cdots,n$ y estas ecuaciones están dadas por la ecuación con vectores columna:

$$A \begin{bmatrix} x_{11} \\ x_{21} \\ \cdots \\ x_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad (3.5)$$

Hay también n ecuaciones que incluyen las n incógnitas x_{i2} para $i = 1,2,\cdots,n$, y así sucesivamente. Además de la ecuación (3.5) hay que resolver:

$$A\begin{bmatrix} x_{12} \\ x_{22} \\ \cdots \\ x_{n2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}, \cdots, A\begin{bmatrix} x_{1n} \\ x_{2n} \\ \cdots \\ x_{nn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{bmatrix}$$

Cada sistema tiene la matriz de coeficientes A, por lo que podemos formar la matriz aumentada:

$$[A|I] = \left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{in} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}\right]$$

Mediante operaciones elementales y usando Gauss-Jordan se obtiene:

$$[I|D] = \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & d_{11} & d_{12} & \cdots & d_{1n} \\ 0 & 1 & \cdots & 0 & d_{21} & d_{22} & \cdots & d_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & d_{n1} & d_{n2} & \cdots & d_{nn} \end{array}\right]$$

De donde $D=A^{-1} = (d_{ij})$.

Ejemplo 3.12

Encuentre la inversa de la matriz A.

$$A=\begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 0 & -1 & 1 \end{bmatrix}$$

Solución

Escribimos $[A|I]$ y efectuamos operaciones elementales para obtener $[I|A^{-1}]$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & -1 & 4 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \approx \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & -5 & -2 & -2 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2/5 & 2/5 & -1/5 & 0 \\ 0 & -1 & 7/5 & 2/5 & -1/5 & 1 \end{array}\right] \approx \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1/7 & 3/7 & -15/7 \\ 0 & 1 & 0 & 2/7 & -1/7 & -2/7 \\ 0 & 0 & 1 & 2/7 & -1/7 & 5/7 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -3/7 & 5/7 & -11/7 \\ 0 & 1 & 0 & 2/7 & -1/7 & -2/7 \\ 0 & 0 & 1 & 2/7 & -1/7 & 5/7 \end{array}\right] = \left[I|A^{-1}\right]$$

De aquí se tiene:

$$\begin{bmatrix} -3/7 & 5/7 & -11/7 \\ 2/7 & -1/7 & -2/7 \\ 2/7 & -1/7 & 5/7 \end{bmatrix} = A^{-1}$$

Dada una matriz, a menudo no se sabe de antemano si es invertible o no. Si se trata de aplicar este procedimiento a una matriz que no es invertible, en algún momento durante el proceso, aparecerá un renglón de ceros en el lado izquierdo, como lo ilustra el siguiente ejemplo:

Ejemplo 3.13

Encuentre la inversa de la matriz:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & -1 & 4 \\ 2 & 4 & 6 \end{bmatrix}$$

Solución

Escribimos $[A|I]$ y efectuamos operaciones elementales para obtener $[I|A^{-1}]$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 3 & -1 & 4 & 0 & 1 & 0 \\ 2 & 4 & 6 & 0 & 0 & 1 \end{array}\right] \approx \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 3 & -1 & 4 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 1 \end{array}\right]$$

De aquí se puede observar que ya no es posible seguir efectuando operaciones elementales, por lo cual la matriz A no tiene inversa.

Podemos resumir el procedimiento para encontrar la inversa de una matriz de la siguiente manera:

Cálculo de A^{-1}:

PASO 1 Formar la matriz aumentada $[A|I]$.

PASO 2 Aplicar el método de Gauss-Jordan para reducir $[A|I]$ a $[I|D]$ Si se puede hacer la reducción entonces $A^{-1}=D$. De no ser así, A^{-1} no existe.

CONDICIONES PARA QUE EXISTA A^{-1}

Teorema 3.3

Si A es una matriz de $n \times n$, entonces las siguientes afirmaciones son equivalentes, es decir, todas son verdaderas o todas son falsas:

a) A es invertible.
b) $AX = 0$ únicamente tiene la solución trivial.
c) A es equivalente por renglones a I_n

Demostración

a) → b)

Suponga que A es invertible siendo x_0 cualquier solución de $AX=0$, por lo tanto $Ax_0=0$. Multiplicando los dos lados de la ecuación por A^{-1} resulta $A^{-1}(Ax_0)=A^{-1}0$, o bien: $(A^{-1}A)x_0=0$, es decir: $Ix_0=0$ de donde $x_0=0$, por lo cual $AX=0$ únicamente tiene la solución trivial.

b) → c)

Sea $AX=0$ la forma matricial asociada al sistema:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{array} \qquad (3.6)$$

Suponga también que el sistema solamente tiene la solución trivial. Si se resuelve empleando la eliminación de Gauss-Jordan, el sistema de ecuaciones correspondiente a la forma escalonada de la matriz aumentada será:

$x_1 = 0$
$x_2 = 0$
……….
$x_n = 0$

Así, la matriz aumentada asociada a (3.6) es igual a:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right]$$

Y se puede llevar a la matriz aumentada

$$\left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right]$$

Mediante una serie de operaciones elementales en los renglones. Si se elimina la última columna (de ceros) en ambas matrices, se puede concluir que A se redujo a I_n mediante una serie de operaciones elementales en los renglones, es decir, A es equivalente por renglones a I_n.

c) $\rightarrow$ a)

Suponga que A es equivalente por renglones a I_n o, lo que es lo mismo, A se puede reducir a I_n mediante una serie finita de operaciones elementales en los renglones, esto se puede hacer multiplicando por la izquierda por una matriz elemental adecuada. De esta manera, es posible hallar matrices elementales E_1, E_2, $\cdots$, E_k tales que:

$$E_k \cdots E_2 E_1 A = I_n$$

Como $E_1, E_2, \cdots, E_k$ son invertibles, pre-multiplicando en forma sucesiva los dos lados de la ecuación por $E_k^{-1}, \cdots, E_2^{-1}, E_1^{-1}$ se obtiene:

$$A = E_1^{-1}\, E_2^{-1} \cdots\, E_k^{-1}\, I_n = E_1^{-1}\, E_2^{-1} \cdots\, E_k^{-1}$$

Como A se puede expresar como un producto de matrices invertibles, entonces A es invertible.

Ejemplo 3.14

Calcule la inversa de la matriz: $A = \begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix}$

Solución

$$\left[\begin{array}{cc|cc} 2 & 9 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right] \approx \underset{E_{12}}{\left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 2 & 9 & 1 & 0 \end{array}\right]} \approx \underset{E_{12}^{(-2)}}{\left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 0 & 1 & 1 & -2 \end{array}\right]} \approx \underset{E_{21}^{(-4)}}{\left[\begin{array}{cc|cc} 1 & 0 & -4 & 9 \\ 0 & 1 & 1 & -2 \end{array}\right]}$$

Por lo tanto:

$$A^{-1} = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix}$$

Ahora expresamos A como producto de matrices elementales:

Si $\quad E_{21}^{(-4)}\, E_{12}^{(-2)}\, E_{12}\; A = I$

Entonces $\quad A = E_{12}\; E_{12}^{(2)}\; E_{21}^{(4)}$

Donde:

$$E_{12} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad E_{12}^{(2)} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad E_{21}^{(4)} = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$$

Recuerde que queremos resolver un sistema $Ax=b$, esto lo podemos hacer a través del método de eliminación de Gauss y obtener un sistema equivalente $Ux=c$, donde U es una matriz triangular superior.

Ejemplo 3.15

Sea el sistema:

$$2x_1 + x_2 + x_3 = 1$$
$$4x_1 + x_2 = -2$$
$$-2x_1 + 2x_2 + x_3 = 7$$

En términos matriciales se tiene:

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 7 \end{bmatrix}$$

Efectuando operaciones elementales se tiene:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 4 & 1 & 0 & -2 \\ -2 & 2 & 1 & 7 \end{array}\right] \underset{E_{12}^{(-2)}}{\approx} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ -2 & 2 & 1 & 7 \end{array}\right] \underset{E_{13}^{(1)}}{\approx} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 3 & 2 & 8 \end{array}\right] \underset{E_{23}^{(3)}}{\approx} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \end{array}\right]$$

Y se obtiene:

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \\ -4 \end{bmatrix}$$

Las matrices elementales usadas son:

$$E_{12}^{(-2)} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_{13}^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad E_{23}^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$$

Todas estas matrices elementales son triangulares inferiores ya que:

a) Si se tiene $E_p^{(a)}$ las matrices son diagonales; es decir, son triangulares inferiores y superiores.
b) Si se tiene $E_{pq}^{(\beta)}$ las matrices son triangulares inferiores si $q>p$.

Por lo tanto, las operaciones matriciales que convierten a A en U son:

$$E_{23}^{(3)}\, E_{13}^{(1)}\, E_{12}^{(-2)} A = U$$

De manera semejante:

$$E_{23}^{(3)}\, E_{13}^{(1)}\, E_{12}^{(-2)}\, b = c$$

3.8. FACTORIZACIÓN LU

La factorización LU de sus siglas en inglés (lower and upper matrix) es una manera de resumir el proceso de eliminación gaussina y nos permite ver el número de operaciones elementales que se efectuaron, por lo que responde a las siguientes preguntas: ¿Qué hacemos para regresar de U a A? ¿Cómo podemos deshacer los pasos de la eliminación gaussiana?

No es difícil deshacer un solo paso por ejemplo; en el paso a), en lugar de restar, sumamos al segundo renglón el doble del primero, de esta manera se invierte la matriz elemental $E_{12}^{(-2)}$ y se tiene:

$$E_{12}^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Análogamente, para b) y c) tenemos las matrices inversas respectivas:

$$E_{13}^{(-1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad E_{23}^{(-3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix}$$

Dichas matrices deshacen por separado cada uno de los pasos a), b) y c). Para deshacer todo el proceso y ver qué matriz regresa a la matriz U en A, nos fijamos en que: c) fue el último paso en realizarse al ir de A a U, por lo que debe ser el primero en invertirse cuando se va en dirección contraria, de aquí que:

$$A = E_{12}^{(2)}\, E_{13}^{(-1)}\, E_{23}^{(-3)}\, U$$

La matriz L que lleva a U de regreso a A debe ser el producto de las 3 matrices:

$$L = E_{12}^{(2)} E_{13}^{(-1)} E_{23}^{(-3)}$$

De aquí que $A = LU$, es la factorización de la matriz A como producto de dos matrices L triangular inferior (L=lower) y U triangular superior (U=upper).

La matriz L es la clave de la eliminación gaussiana, es el vínculo entre la matriz A y la matriz U a la que llegamos.

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix}$$

La matriz L así construida es triangular inferior con unos en la diagonal principal. Los elementos debajo de la diagonal son precisamente los multiplicadores 2, −1, −3, usados en los 3 pasos de la eliminación. También podemos observar que:

$$L^{-1} = E_{23}^{(3)} E_{13}^{(1)} E_{12}^{(-2)}$$

Por lo tanto:

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{bmatrix}$$

Ejemplo 3.16

Reduzca la matriz A, a la forma triangular superior U y cree la matriz L.

$$A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 8 & 4 \\ -1 & 3 & 4 \end{bmatrix}$$

Solución

Para hacer más explícita la forma en que se construye la matriz L, haremos el proceso en dos columnas: en una de ellas se va a iniciar con la matriz A y en la otra con la matriz identidad; se van a ir efectuando las operaciones elementales simultáneamente en ambas matrices para así, al finalizar el proceso, obtener las matrices U y L.

Reducción de A en U **Creación de L**

$$A=\begin{bmatrix} 1 & 3 & -1 \\ 2 & 8 & 4 \\ -1 & 3 & 4 \end{bmatrix} \qquad I=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_{12}^{(-2)} \qquad E_{12}^{(2)}$$

$$\approx \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ -1 & 3 & -15 \end{bmatrix} \qquad \approx \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_{13}^{(1)} \qquad E_{13}^{(-1)}$$

$$\approx \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ 0 & 6 & 3 \end{bmatrix} \qquad \approx \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$E_{23}^{(-3)} \qquad E_{23}^{(3)}$$

$$\approx \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ 0 & 0 & -15 \end{bmatrix} \qquad \approx \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 3 & 1 \end{bmatrix}$$

Ahora veremos cómo se puede utilizar la información conservada en L de este ejemplo para resolver un sistema lineal $Ax=b$ con esta matriz A.

Ejemplo 3.17

Utilice la matriz L para hallar el vector columna c que se presentaría si redujéramos (A/b) en (U/c) aplicando esas mismas operaciones en los renglones.

$$b=\begin{bmatrix} -4 \\ 2 \\ 4 \end{bmatrix}$$

$I_{21}=2$ Significa multiplicar el renglón 1 por −2 y sumarlo al renglón 2 $\approx \begin{bmatrix} -4 \\ 10 \\ 4 \end{bmatrix}$

$I_{31}=-1$ Significa multiplicar el renglón 1 por 1 y sumarlo al renglón 3 $\approx \begin{bmatrix} -4 \\ 10 \\ 0 \end{bmatrix}$

$I_{32}=3$ Significa multiplicar el renglón 2 por 3 y sumarlo al renglón 3 $\approx \begin{bmatrix} -4 \\ 10 \\ 30 \end{bmatrix} = c$

PROPIEDADES

1. La matriz A puede escribirse, mientras que ninguno de los pivotes sea cero, como un producto LU de una matriz triangular inferior L y una matriz triangular superior U.
2. Los elementos de la diagonal principal de L son unos, y debajo de la diagonal están los multiplicadores I_{ij}.
3. U es la matriz de coeficientes que obtenemos después de la eliminación y antes de la sustitución regresiva, los elementos de su diagonal son los pivotes.

Ejemplo 3.18

Considere la matriz:

$$A = \begin{bmatrix} 6 & 2 & 1 & -1 \\ 2 & 4 & 1 & 0 \\ 1 & 1 & 4 & -1 \\ 1 & 0 & -1 & 3 \end{bmatrix}$$

Entonces una factorización de $A = LU$ sería:

$$L = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \qquad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}$$

Los 16 elementos de A se pueden usar para determinar parcialmente los 10 elementos desconocidos de L y el mismo número que los de U. Sin embargo, si el procedimiento nos debe llevar a una solución única, se necesitan cuatro

condiciones adicionales para los elementos de L y U. El método que hemos usado consiste en que $l_{11} = l_{22} = l_{33} = l_{44} = 1$; este método se conoce como **Método de Doolittle.** Para este ejemplo se tiene:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/3 & 1 & 0 & 0 \\ 1/6 & 1/5 & 1 & 0 \\ -1/6 & 1/10 & -9/37 & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 6 & 2 & 1 & -1 \\ 0 & 10/3 & 2/3 & 1/3 \\ 0 & 0 & 37/10 & -9/10 \\ 0 & 0 & 0 & 191/74 \end{bmatrix}$$

Otro método comúnmente usado para factorizar es el **Método de Crout**, el cual requiere que los elementos de la diagonal principal de U sean iguales a uno, y que la matriz A sea positiva definida o estrictamente dominante diagonalmente.

Teorema 3.4

Si el procedimiento de eliminación gaussiana puede aplicarse al sistema $Ax=b$ sin intercambio de renglones, entonces la matriz A puede factorizarse como producto de una matriz triangular inferior L con una matriz triangular superior U.

$$A = LU$$

Ventajas de la forma *LU*

1. Se puede ver que la forma LU sirve para ahorrar trabajo si se tiene una matriz de coeficientes A y se modifica el vector b por un nuevo vector b', entonces ya no se tiene que volver a resolver todo el sistema.
2. Mayor rapidez en tiempo de computadora. Supongamos que a cada división y cada multiplicación-sustracción le asignamos una sola operación. Al principio, cuando la primera ecuación tiene longitud n, lleva n operaciones por cada cero que logramos en la primera columna: una para encontrar el múltiplo l y las otras para encontrar las nuevas entradas a lo largo del renglón. Hay $n-1$ renglones bajo la primera, de modo que el primer paso de la eliminación necesita $n(n-1) = n^2-n$ operaciones. (Otra forma de obtener n^2-n es la siguiente: es necesario cambiar las n^2 entradas exceptuando la primera fila). Observe ahora que las etapas posteriores son más rápidas, porque las ecuaciones se vuelven progresivamente más cortas: cuando la eliminación ha llegado a k ecuaciones solo se necesitan k^2-k operaciones para desalojar

la columna debajo del pivote, empleando el mismo razonamiento aplicado a la primera etapa, cuando k era igual a n, esto es:

$$P = (1^2 + \cdots + n^2) - (1 + \cdots + n) = \frac{n(n+1)(2n+1)}{6} = \frac{-n(n+1)}{2} = \frac{n^3 - n}{3}$$

Para n suficientemente grande, $P \approx \frac{n^3}{3}$

La sustitución regresiva es más rápida. La última incógnita se encuentra en una operación (una división entre el último pivote), la penúltima requiere dos, y así sucesivamente. La sustitución regresiva requiere un total de operaciones de:

$$O = 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \approx \frac{n^2}{2} \text{ para la } n \text{ grande.}$$

Una vez que conocemos LU podemos encontrar la solución x' correspondiente a cualquier nuevo lado derecho b', en solo n^2 operaciones. (Usar L^{-1} o U^{-1} solo nos haría perder tiempo, igual si se tiene A^{-1}, aunque el producto $A^{-1}b$ también requiere n^2 operaciones). Utilizar la técnica LU tiene al menos dos ventajas: En primer lugar, para hallar U se requieren $n^3/3$ operaciones para n grande, y para hallar A^{-1} se necesitan n^3 operaciones; si $n=1000$, la diferencia en tiempo de computadora es grande. Además lo que queremos es la solución, no todas las entradas de la inversa.

Observaciones a la forma *LU*

1. La forma LU es "asimétrica" en un sentido: a lo largo de su diagonal principal U contiene los pivotes, mientras que L siempre tiene unos. Esto es fácil de corregir: basta factorizar de U una matriz diagonal D constituida por los pivotes $d_1, d_2, \cdots, d_n$:

$$U = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \cdots & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & u_{12}/d_1 & u_{13}/d_1 & \cdots & u_{1n}/d_1 \\ & 1 & u_{23}/d_1 & \cdots & u_{2n}/d_1 \\ & & & \cdots & \cdots \\ & & & & 1 \end{bmatrix}$$

Y se escribe A como $A = LDU$. L es triangular inferior con unos en la diagonal, U es triangular superior con unos en la diagonal y D es la matriz diagonal de los pivotes.

2. No hay mucha libertad al efectuar el proceso de eliminación, ya que las operaciones en los renglones de manera aleatoria podrían fácilmente destruir en un paso los ceros generados en un paso anterior.

Ejemplo 3.19

Se expresa A en la factorización LDU como sigue:

$$A = \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{bmatrix}} = \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix}} \underset{D}{\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{bmatrix}} \underset{U}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}$$

3.9. LA FORMA *PA*=*LU*

Existe la variante de que una matriz no se pueda escribir directamente en la forma LU por ejemplo, ¿Qué pasa cuando el número que se va a usar como pivote es igual a cero?

$$\begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

No podemos usar ningún múltiplo de la primer ecuación para aniquilar el coeficiente 3. Entonces intercambiamos las dos ecuaciones y en este caso sencillo la matriz será triangular superior y el sistema:

$$3x_1 + 4x_2 = b_2$$
$$2x_2 = b_1$$

Se puede resolver fácilmente por sustitución regresiva. Expresando esto en términos matriciales, necesitamos encontrar la matriz de permutación que produce el intercambio de renglones, esto es:

Entonces:

$$PA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}$$

Lo mismo sucede con b por lo que $PAx = Pb$.

Sí A es una matriz más grande; por ejemplo, de 4×4 y con $a_{22} = 0$

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & d & 6 \\ 0 & c & 7 & 8 \end{bmatrix}$$

Nos fijamos en la columna con pivote igual a cero; si hay algún elemento diferente de cero, se intercambia el renglón, si no, entonces la matriz es singular. En este caso se intercambian las filas 2 y 4; la matriz de permutación P_{24} que produce el intercambio es:

$$P_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Note que P_{24} es la matriz I_4 con las filas 2 y 4 intercambiadas. En general, la matriz de permutación P_{kl} es la idéntica con las filas k y l intercambiadas y $P_{kl}A$ produce el mismo intercambio de filas en A, esto es:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & d & 6 \\ 0 & c & 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & c & 7 & 8 \\ 0 & 0 & d & 6 \\ 0 & 0 & 5 & 6 \end{bmatrix}$$

Continuamos con la siguiente columna: si $d = 0$ hacemos P_{34} y observamos que el sistema es no singular, entonces tenemos $P_{34}P_{24}A$ y terminamos. Otros dos ejemplos serían:

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 6 & 9 \\ 2 & 8 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 2 & 4 \end{bmatrix} \rightarrow \overset{PA}{\begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 5 \end{bmatrix}} \text{ (no singular)}$$

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 6 & 9 \\ 3 & 9 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 0 & 2 \end{bmatrix} \text{ (singular)}$$

¿Qué sucede con la factorización LU (o LDU) cuando hay intercambio de renglones? Se obtiene U triangular superior, pero ahora no solo tenemos matrices elementales E_{ij} sino también matrices de permutación P_{kl}, cuyo producto no es triangular inferior. Así que A no puede expresarse como $A=LU$. Sin embargo, podemos reemplazar A por la matriz PA. Donde P es una matriz de permutación que es el producto de cada una de las permutaciones, resumiendo:

Observación:
En el caso no singular existe una matriz de permutación P que reordena los renglones de A de modo que PA admite una factorización con pivotes diferentes de cero, y $PA=LU$ (o LDU). En éste caso hay solución única a $Ax=b$, la cual se encuentra a través de la eliminación. En el caso singular ningún reordenamiento puede producir pivotes distintos de cero.

Ejemplo 3.20
Dada la matriz A, factorícela a la forma LU

$$A=\begin{bmatrix} 1 & 3 & 2 \\ -2 & -6 & 1 \\ 2 & 5 & 7 \end{bmatrix}$$

Solución

$$\begin{bmatrix} 1 & 3 & 2 \\ -2 & -6 & 1 \\ 2 & 5 & 7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & -1 & 3 \end{bmatrix}$$

Se observa que se necesitan intercambiar las filas 2 y 3; así podemos usar:

$$P_{23}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

De tal forma que:

$$PA=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 3 & 2 \\ -2 & -6 & 1 \\ 2 & 5 & 7 \end{bmatrix}$$

De donde:

$$L=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

Y se verifica que $PA = LU$

$$PA=\begin{bmatrix} 1 & 3 & 2 \\ 2 & 5 & 7 \\ -2 & -6 & 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 5 \end{bmatrix}=LU$$

La factorización LU también se puede usar para resolver sistemas donde la matriz A está elevada a una potencia, como se muestra en el siguiente ejemplo:

Ejemplo 3.21

Resuelva el sistema lineal $A^3x=b$.

$$A=\begin{bmatrix} 1 & 2 & -1 \\ -2 & -5 & 3 \\ -1 & -3 & 0 \end{bmatrix} \qquad b=\begin{bmatrix} 9 \\ -17 \\ -44 \end{bmatrix}$$

Solución

$A^3x = b$	lo podemos escribir como:
$A(A^2x) = b$	sea $y = A^2x$ entonces nos queda
$Ay = b$	como $A^2x = y$ entonces $A(Ax) = y$ entonces sea $z = Ax$ y obtenemos
$Az = y$	se resuelve $Az = y$ para z y finalmente resolvemos
$Ax = z$	para la x deseada.

Como en todo el proceso siempre usamos la misma matriz de coeficientes A, es rentable hallar L y U.

$$U=\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \qquad L=\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

Resolvemos para $Ay = b$ con la información conservada en L y se tiene

$$b=\begin{bmatrix} 9 \\ -17 \\ -44 \end{bmatrix} \cong \begin{bmatrix} 9 \\ 1 \\ -35 \end{bmatrix} \cong \begin{bmatrix} 9 \\ 1 \\ -36 \end{bmatrix}$$

Entonces, usando U tenemos

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 1 \\ -36 \end{bmatrix}$$

Y de aquí

$$y=\begin{bmatrix} -7 \\ 17 \\ 18 \end{bmatrix}$$

Para resolver $Az = y$ se aplica la información conservada en L a y

$$y=\begin{bmatrix} -7 \\ 17 \\ 18 \end{bmatrix} \cong \begin{bmatrix} -7 \\ 3 \\ 8 \end{bmatrix} \cong \begin{bmatrix} -7 \\ 3 \\ 8 \end{bmatrix}$$

Nuevamente usamos U

$$\begin{bmatrix} 1 & 2 & -2 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} -7 \\ 3 \\ 8 \end{bmatrix}$$

Y de aquí

$$z=\begin{bmatrix} 3 \\ -7 \\ -4 \end{bmatrix}$$

Para resolver $Ax=z$ usamos la información conservada en L para aplicarla a z y obtenemos

$$z=\begin{bmatrix} 3 \\ -7 \\ -4 \end{bmatrix} \approx \begin{bmatrix} 3 \\ -1 \\ -4 \end{bmatrix} \approx \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix} \approx \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$$

Nuevamente usamos U

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$$

Y se obtiene:

$$x=\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

3.10. MATRICES ESPECIALES

Algunas matrices por tener características muy bien determinadas se consideran en esta sección como matrices especiales tales son los casos de las matrices estrictamente dominantes diagonalmente que es importante en el caso de resolver sistemas de ecuaciones grandes (mayores a 100) a través de métodos como el de Jacobi. Otro es el caso de las matrices positivas definidas que siempre tienen inversa.

Definición 3.4

Se dice que la matriz A de $n \times n$ es estrictamente dominante diagonalmente si satisface:

$$|a_{ii}| > \sum_{\substack{j=i \\ j \neq i}}^{n} |a_{ij}| \qquad \text{Para toda } i=1, \cdots, n$$

Y solo dominante diagonalmente sí

$$|a_{ii}| \geq \sum_{\substack{j=i \\ j \neq i}}^{n} |a_{ij}| \qquad \text{para toda } i=1, \cdots, n$$

Ejemplo 3.22

Sea la matriz A como sigue:

$$\begin{bmatrix} 7 & 2 & 0 \\ 3 & 5 & -1 \\ 0 & 5 & -6 \end{bmatrix}$$

Se dice que A es estrictamente dominante diagonalmente ya que

$$|7| > |2| + |0| \qquad |5| > |3| + |-1| \qquad |-6| > |0| + |5|$$

Teorema 3.5

Si A es una matriz de $n \times n$ estrictamente dominante diagonalmente, entonces A es no-singular. Además, se puede efectuar eliminación gaussiana en cualquier sistema lineal de la forma $Ax=b$ para obtener su solución única sin intercambio de renglones o columnas.

Demostración

Para probar que A es no singular, sea el sistema lineal $Ax=0$, y supongamos que existe una solución $x=(x_i) \neq 0$ a este sistema, en este caso para alguna k tal que:

$$0 < |x_k| = \max_{l \leq j \leq n} |x_j|$$

Como $\sum_{j=1}^{n} a_{ij} \ x_j = 0$ para toda $i = 1, 2, \cdots, n$ es decir

$$
\begin{aligned}
&a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n = 0\\
&a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n = 0\\
&\vdots\\
&a_{n1}\,x_1 + a_{n2}\,x_2 + \cdots + a_{nn}\,x_n = 0
\end{aligned}
$$

Para $i=k$

$$a_{k1}\,x_1 + a_{k2}\,x_2 + \cdots + a_{kk}\,x_k + \cdots + a_{kn}\,x_n = 0$$

Despejando se tiene

$$a_{kk}\,x_k - \sum_{\substack{j=i\\ j\neq k}}^{n} a_{kj} \quad x_j$$

Esto implica que

$$|a_{kk}|\,|x_k| \ \leq -\sum_{\substack{j=i\\ j\neq k}}^{n} |\,a_{kj}\,|\,|x_j|$$

O bien

$$|a_{kk}| \leq \sum_{\substack{j=i\\ j\neq k}}^{n} |\,a_{kj}\,| \frac{|\,x_j\,|}{|\,x_k\,|} \leq \sum_{j=1}^{n} a_{kj}$$

Lo que contradice el hecho de que A sea estrictamente dominante diagonalmente, y esto implica que la única solución de $Ax = 0$ es $x=0$ que como ya habíamos visto en el teorema 3.3, es equivalente a la no singularidad de A.

> ***Definición 3.5***
> Sea A una matriz simétrica, A es positiva definida si $x^t Ax > 0$ para toda $x \neq 0$

Ejemplo 3.23

Verifique si las siguientes matrices son positivas definidas

$$A=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \qquad B=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \qquad C=\begin{bmatrix}2 & -1 & 0\\-1 & 2 & -1\\0 & -1 & 2\end{bmatrix}$$

Solución

$$x^tAx=[x_1,x_2]\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=x_1^2+2x_2^2>0$$

Para toda x, por lo tanto A es positiva definida. De igual forma se efectúa para B y se tiene:

$$x^tBx=4{x_1}^2-{x_2}^2$$

Si $x_1=1$ y $x_2=3$ entonces $x^tBx<0$ por lo tanto B no es positiva definida.

Para C se tiene:

$$x^tCx={x_1}^2+(x_1-x_2)^2+(x_2-x_3)^2+{x_3}^2>0$$

A menos que $x_1=x_2=x_3=0$.

TEOREMA 3.6

Sí A es una matriz de $n\times n$ positiva definida, entonces A es no singular. Además, la eliminación gaussiana se puede aplicar a cualquier sistema lineal de la forma $Ax=b$ para obtener su solución única sin intercambios de renglones o de columnas.

Demostración

Si x≠0 es un vector que satisface $Ax=0$, entonces $x^tAx=0$ es imposible ya que A es positiva definida, lo que implica que $Ax=0$ solo tiene la solución trivial, por lo tanto, es no singular. La otra parte de la demostración se omite.

3.11. RESUMEN

1. Un sistema lineal tiene asociada una matriz partida (o aumentada), que tiene la matriz de coeficientes del sistema a la izquierda de la partición y el vector columna de constantes a la derecha de la partición.

2. Las operaciones elementales en las filas son:
   » Intercambio de filas.
   » Multiplicación de una fila por una constante distinta de cero.
   » Multiplicación de una fila por una constante distinta de cero y sumar a otra fila.
3. En el método de Gauss, se resuelve un sistema lineal cuadrado usando operaciones elementales en filas y reduciendo la matriz partida de modo que la parte izquierda de la partición se convierta en triangular superior. Luego se halla la solución mediante sustitución regresiva.
4. El método de Gauss-Jordan es análogo al método de Gauss, excepto que la matriz de coeficientes del sistema se reduce a la matriz identidad I. Entonces, la solución aparece como el vector columna a la derecha de la partición en la matriz aumentada reducida.
5. Se pueden resolver al mismo tiempo varios sistemas lineales con la misma matriz de coeficientes aumentando la matriz de coeficientes con los vectores columna de las constantes de todos los sistemas.
6. Sea A una matriz cuadrada. Una matriz cuadrada C tal que $CA=AC=I$ es la inversa de A y se expresa $C=A^{-1}$. Si tal inversa A^{-1} de A existe, entonces se dice que A es invertible. La inversa de una matriz invertible A es única. Una matriz cuadrada que no tiene inversa se llama singular.
7. La inversa de una matriz cuadrada A existe si, y solo si, A puede reducirse a la matriz identidad I mediante operaciones elementales en las filas, o también, si y solo si, A es igual a un producto de matrices elementales. En este caso, A es igual al producto, tomado de izquierda a derecha, de las inversas de las matrices elementales sucesivas que corresponden a la sucesión de operaciones en las filas que reduce A en I.
8. Para hallar A^{-1}, si existe, se forma la matriz partida $(A^{3}I)$ y se aplica el método de Gauss–Jordán para reducir esta matriz a $(I^{3}D)$. Si se puede hacer esto, entonces $A^{-1}=D$. De lo contrario, A no es invertible.
9. La inversa de un producto de matrices invertibles es el producto de las inversas en orden contrario.
10. Si A es una matriz invertible de $n\times n$ que puede reducirse por filas a una matriz triangular superior U sin intercambio de filas, entonces existe una matriz triangular inferior L de $n\times n$ tal que $A=LU$.
11. La matriz L se puede hallar como sigue: se comienza con la matriz identidad I de $n\times n$. Si durante la reducción de A en U se multiplica la fila i por r y se suma el resultado a la fila k, reemplazar el cero de la fila k y columna i de la

matriz identidad por $-r$. El resultado final obtenido de la matriz identidad es la matriz L.

12. Una vez que A se ha reducido a U y se ha hallado L se puede encontrar la solución de $Ax=b$ para cualquier nuevo vector columna b.
13. Si A es como en el punto (10), entonces A tiene una factorización única de la forma $A=LDU$, donde:
    L es triangular inferior con todas las entradas diagonales 1,
    U es triangular superior con todas las entradas diagonales 1,
    D es una matriz diagonal con todas las entradas en la diagonal igual a los pivotes.
14. Para cualquier matriz invertible A, existe una matriz de permutación P tal que PA se puede reducir por filas a una matriz triangular superior U y tiene las propiedades descritas para A en los puntos 10, 11, 12 y 13.

3.12. APLICACIONES DE SISTEMAS DE ECUACIONES

Los sistemas de ecuaciones lineales tienen infinidad de aplicaciones, en todas las ramas de la ingeniería, en medicina, biología, economía, finanzas por mencionar algunas. Por cuestiones de espacio y considerando el enfoque de esta publicación y a quienes va dirigida, se presentan las siguientes.

CADENAS DE MARKOV

(Tomado de *Aplicaciones de álgebra lineal*, Rorres y Anton, 1977)

En las cadenas de Markov podemos ver el caso de ecuaciones lineales que se describen a través de matrices y que tienen probabilidades asociadas como elementos o entradas, se va a operar con estas matrices para representar diferentes estados como se explica a continuación.

Supongamos un sistema, físico o matemático de naturaleza tal que, en cualquier momento, presenta uno de los estados posibles de un número finito de estados. Por ejemplo, el estado atmosférico en una cierta ciudad puede ser uno de los siguientes tres estados posibles: sol brillante, nublado o lluvioso. Un individuo puede estar en uno de los cuatro estados emocionales posibles: contento, triste, enojado o aprensivo. Suponga que uno de estos sistemas pasa, con el tiempo, de llevar un registro de los estados del sistema correspondientes a estos tiempos. Si

se encuentra que la transición de un estado a otro no está predeterminada, sino que más bien ocurre en función de ciertas probabilidades que dependen de la historia del sistema, el proceso recibe entonces el nombre de proceso estocástico. Si, además, estas probabilidades de transición dependen solamente de la historia inmediata del sistema, es decir, si el estado del sistema en una observación cualquiera depende solo de su estado en la observación inmediata anterior, el proceso recibe el nombre de proceso de Markov o cadena de Markov.

Suponga que el sistema en observación tiene k estados posibles, a los cuales se les identifica con 1, 2, $\cdots$, k. Para describir las transiciones de un estado a otro, se darán las siguientes definiciones:

Definición 3.6

La probabilidad de transición P_{ij} $(i,j=1,2,\cdots,k)$ es la probabilidad de que ocurra que, si el sistema está en el estado j en una observación cualquiera, esté en el estado i en la siguiente observación.

Por ejemplo, si el estado 2 corresponde a un día lluvioso en Detroit y el estado 3 a un día nublado, P_{32} es la probabilidad de que cambie el tiempo en Detroit de lluvioso a nublado en dos días consecutivos.

Como son probabilidades, los números p_{ij} deben estar comprendidos en el intervalo [0, 2]. Además, para cualquier valor fijo de j=1, 2,$\cdots$, k, deben tenerse que:

$$p_{1j} + p_{2j} + \cdots + p_{kj} = 1 \qquad (3.7)$$

Expresión que dice que si el sistema, en una observación, está en el estado j estará, con toda seguridad, en uno de los k estados cuando se haga la siguiente observación.

Con estas k^2 probabilidades de transición se puede formar una matriz de , llamada matriz de transición del proceso de Markov. La ecuación (3.7) expresa también que la suma de las entradas, en cada columna de P, es igual a uno. En forma más general, se puede dar la siguiente definición:

Definición 3.7

Una matriz de transición $P=[p_{ij}]$ es una matriz cuadrada, con entradas no negativas, en la que la suma de cada columna es igual a la unidad.

Entonces, todo proceso de Markov determina una matriz de transición. A las matrices de transición se les llama también matrices de Markov, matrices de probabilidad o matrices estocásticas.

Ejemplo 3.24

Revisando los registros de donativos, la oficina de la sociedad de alumnos de un colegio encuentra que el 80% de los alumnos que contribuyeron a la integración del fondo anual, en un cierto año, contribuirán también al año siguiente y que el 30% de los que no contribuyeron en ese año, también contribuirán al siguiente. Esta situación se puede describir como un proceso de Markov con dos estados: el estado 1 corresponde a un alumno que da un donativo en un año cualquiera y el estado 2 a un alumno que no da donativo en ese mismo año. La matriz de transición del proceso es:

$$p=\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$$

El objetivo de esta sección es predecir el estado de un sistema correspondiente a un proceso de Markov. Sin embargo, como el sistema no es determinista, las predicciones se harán en función de las probabilidades. Así, es necesarito introducir la siguiente definición:

> ***Definición 3.8***
> Un vector de probabilidad es un vector columna, con entradas no negativas, en el que la suma de sus elementos es igual a la unidad.

Y se puede pronosticar el resultado de una observación futura de un sistema que sea un proceso de Markov, en función de un vector de probabilidad, si aplicamos la siguiente definición:

> ***Definición 3.9***
> Se dice que los vectores de probabilidad para son los vectores de estado de un proceso de Markov, si la componente de orden i, $x_i^{(n)}$ de $x^{(0)}$, es la probabilidad de que el sistema esté en el estado i cuando se hace la observación en n.

En particular, al vector $x^{(0)}$ se llama vector del estado inicial del proceso de Markov. Aplicando el siguiente teorema, se verá que el vector del estado inicial

es suficiente para determinar todos los vectores de estado correspondientes a observaciones futuras.

Teorema 3.7

Si P es la matriz de transición en un proceso de Markov y $x^{(0)}$ es el vector de estado de la observación n, se tendrá que $x^{(n+1)}= Px^{n}$.

La demostración de este teorema se basa en la teoría de las probabilidades y no se hará aquí. De este teorema se deduce que:

$$\begin{aligned}
x^{(1)} &= Px^{(0)} & & \\
x^{(2)} &= Px^{(1)} &=& P^2\,x^{(0)} \\
x^{(3)} &= Px^{(2)} &=& P^3\,x^{(0)} \\
&\vdots & & \\
x^{(n)} &= Px^{(n-1)} &=& P^n\,x^{(0)}
\end{aligned}$$

Así, el vector de estado inicial $x^{(0)}$ y la matriz de transición P determinan a $x^{(n)}$ para n=1, 2, ⋯

Ejemplo 3.24 (ya presentando con anterioridad)

La matriz de transición del ejemplo 3.23 era:

$$p = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$$

Puede obtenerse un registro de los probables donativos de un nuevo graduado que no dio su donativo en el año que siguió a su graduación. Con este planteamiento, tenemos la seguridad de que el sistema está inicialmente en el estado 2 por lo que el vector del estado inicial es:

$$x^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Y por el teorema 3.7 se tiene:

$$x^{(1)} = Px^{(0)} = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} .3 \\ .7 \end{bmatrix}$$

$$x^{(2)} = Px^{(1)} = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \begin{bmatrix} .3 \\ .7 \end{bmatrix} = \begin{bmatrix} .45 \\ .55 \end{bmatrix}$$

$$x^{(3)} = Px^{(2)} = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \begin{bmatrix} .45 \\ .55 \end{bmatrix} = \begin{bmatrix} .525 \\ .475 \end{bmatrix}$$

Así, después de tres años, se puede esperar que el alumno haga un donativo con una probabilidad de 0.525. Más allá de los tres años, se obtienen los siguientes vectores de estado (con tres decimales):

$$x^{(4)} = \begin{bmatrix} .563 \\ .438 \end{bmatrix}, \quad x^{(5)} = \begin{bmatrix} .581 \\ .419 \end{bmatrix}, \quad x^{(6)} = \begin{bmatrix} .591 \\ .409 \end{bmatrix}, \quad x^{(7)} = \begin{bmatrix} .595 \\ .405 \end{bmatrix}$$

$$x^{(8)} = \begin{bmatrix} .598 \\ .402 \end{bmatrix}, \quad x^{(9)} = \begin{bmatrix} .599 \\ .401 \end{bmatrix}, \quad x^{(10)} = \begin{bmatrix} .599 \\ .401 \end{bmatrix}, \quad x^{(11)} = \begin{bmatrix} .600 \\ .400 \end{bmatrix}$$

Para todos los valores de n mayores que once, se tiene $x(n) = \begin{bmatrix} .600 \\ .400 \end{bmatrix}$

En otras palabras: los vectores de estado tienden a un vector fijo si se aumenta el número de observaciones.

Ejemplo 3.25

Una agencia que renta automóviles tiene tres oficinas de renta que se identifican con 1, 2 y 3 y que están ubicados en ciudades diferentes. Un cliente puede rentar un auto en cualquiera de las tres oficinas y devolverlo en cualquiera de las tres ciudades. El gerente determina que los clientes devuelven los automóviles de acuerdo con las siguientes probabilidades:

Rentado en la oficina

$$\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} .8 & .3 & .2 \\ .1 & .2 & .6 \\ .1 & .5 & .2 \end{bmatrix} \end{array} \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}$$

Devuelto a la oficina

Esta es la matriz de transición del sistema, considerado como un proceso de Markov. La componente de orden i de un vector de estado es la probabilidad de que un auto sea devuelto a la oficina i. En la siguiente tabla se da una lista de vectores de estado $x^{(n)}$ para n= 0, 1, ⋯, 11, para el caso en el que, inicialmente, el auto se renta en la oficina 2.

Para todos los valores de n mayores que once, los vectores de estado son iguales a $x^{(11)}$ considerando tres decimales.

Tabla 3.1 Probabilidades asociadas a los vectores de estado

n	0	1	2	3	4	5	6	7	8	9	10	11
$x_1^{(n)}$	0	.300	.400	.477	.511	.533	.544	.550	.553	.555	.556	.557
$x_2^{(n)}$	1	.200	.370	.252	.261	.240	.238	.233	.232	.231	.230	.230
$x_3^{(n)}$	0	.500	.230	.271	.228	.227	.219	.217	.215	.214	.214	.213

En este ejemplo, deben observarse dos cosas: que no fue necesario saber cuánto tiempo conservaba el cliente el auto en su poder, es decir, un proceso de Markov no requiere intervalos constantes de tiempo entre las observaciones y que los vectores de estado se aproximan o tienden a un vector fijo cuando aumenta n, como ocurrió en el primer ejemplo.

Ejemplo 3.26

A un policía de tránsito se le asigna la labor de controlar el tráfico en los ocho cruceros marcados en la Figura 3.1. Se le instruye en el sentido que permanezca una hora en cada crucero y una vez cubierto el primero, inicie el recorrido quedándose en el mismo crucero o yendo a uno de los inmediatos. Para evitar el establecimiento de un patrón de movimientos, se le dice que escoja al azar un nuevo crucero, con lo cual todas las posibles elecciones tienen la misma probabilidad. Por ejemplo, si el policía se encuentra en el crucero 5, el siguiente puede ser 2, 4, 5 u 8, teniendo cada uno la probabilidad de ¼.

Figura 3.1 Cruceros asignados al policía

Cada día, el policía inicia el crucero en el que estuvo al final del día anterior, considerando al problema como un proceso de Markov, se puede obtener la siguiente matriz de transición:

1	2	3	4	5	6	7	8	1
⅓	⅓	0	⅕	0	0	0	0	2
⅓	⅓	0	0	¼	0	0	0	3
0	0	⅓	⅕	0	⅓	0	0	4
⅓	0	⅓	⅕	¼	0	¼	0	5
0	⅓	0	⅕	¼	0	0	⅓	6
0	0	⅓	0	0	⅓	¼	0	7
0	0	0	0	⅕	0	¼	⅓	8
0	0	0	0	0	¼	¼	⅓	9

Si el policía inicia en la intersección 5, las probables posiciones, hora por hora, están dados por los siguientes vectores de estado presentados en forma tabular.

n	0	1	2	3	4	5	10	15	20	22
$x_1^{(n)}$	0	.000	.133	.116	.130	.123	.113	.109	.108	.107
$x_2^{(n)}$	0	.250	.146	.163	.140	.138	.115	.109	.108	.107
$x_3^{(n)}$	0	.000	.050	.039	067	.073	.100	.106	.107	.107
$x_4^{(n)}$	0	.250	.113	.187	.162	.178	.178	.179	.179	.179
$x_5^{(n)}$	1	.250	.279	.190	.190	.168	.149	.144	.143	.143
$x_6^{(n)}$	0	.000	.000	.050	.056	.074	.099	.105	.107	.107
$x_7^{(n)}$	0	.000	.133	.104	.131	.125	.138	.141	.143	.143
$x_8^{(n)}$	0	.250	.146	.152	.124	.121	.108	.107	.107	.107

Para todos los valores de n mayores que 22, los vectores de estado son iguales a $x^{(22)}$ considerando tres decimales. Por lo tanto, como ocurrió en los dos primeros ejemplos, los vectores de estado tienden a un vector fijo cuando aumenta n.

EL COMPORTAMIENTO DE LOS VECTORES DE ESTADO EN EL LÍMITE

En los tres ejemplos anteriores se vio que los vectores de estado tienden a un cierto vector fijo cuando aumenta el número de observaciones. Cabe ahora preguntar si en un proceso de Markov, los vectores de estado tienden siempre a un vector fijo. Un ejemplo simple demuestra que esto no es cierto:

Ejemplo 3.27

Sean $P=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ y $x^{(0)}=\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Como $P^2 = I$ y $P^3 = P$, se tiene que:

$$x^{(0)} = x^{(2)} = x^{(4)} = \cdots = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Y que

$$x^{(1)} = x^{(3)} = x^{(6)} = \cdots = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Este sistema oscila indefinidamente entre los vectores de estado $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, por lo que no tienden a ningún vector fijo.

Sin embargo, con una ligera condición impuesta a la matriz de transición, se puede demostrar que los vectores de estado sí tienen como límite un vector fijo. Esta condición está dada por la siguiente definición:

> ***Definición 3.10***
> Una matriz de transición es regular si una potencia entera de la misma tiene todas sus entradas positivas.

Así, dada una matriz regular de transición P, existen un entero positivo m tal que todas las entradas de P^m sean positivas. Esto ocurre con las matrices de transición de los ejemplos 3.28 y 3.29 cuando $m=1$. En el ejemplo 3.30, P^4 es la matriz que tiene todas las entradas positivas. En consecuencia, los tres ejemplos tienen una matriz de transición que es regular.

Cuando un proceso de Markov se rige por una matriz de transición regular, se dice que se trata de un proceso regular de Markov. Se verá que todo proceso regular tiene un vector de estado fijo q tal que $P^n x^{(0)}$ tiende a q cuando aumenta

n, independientemente de la selección de $x^{(0)}$. Este resultado es de gran importancia para la teoría de las cadenas de Markov y se basa en el siguiente teorema.

Teorema 3.8

Si P es una matriz regular de transición, ocurre que cuando $n \to \infty$

$$p_n \to \begin{matrix} q_1 & q_1 & \cdots & q_1 \\ q_2 & q_2 & \cdots & q_2 \\ \vdots & \vdots & & \vdots \\ q_k & q_k & \cdots & q_k \end{matrix}$$

Donde los términos q_i son números positivos tales que:

$$q_1 + q_2 + \cdots + q_k = 1$$

Este teorema no se demostrará. Al lector interesado en la demostración se le sugiere que consulte un texto más especializado, como por ejemplo el libro "Finite Markov Chains" por J. Kemeny y J. Snell, editado por Springer-Verlag, New York, Inc., 1976.

Si establece que:

$$Q = \begin{bmatrix} q_1 & q_1 & \cdots & q_1 \\ q_2 & q_2 & \cdots & q_2 \\ \vdots & \vdots & & \vdots \\ q_k & q_k & \cdots & q_k \end{bmatrix} \quad \text{y} \quad q = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_3 \end{bmatrix}$$

Q será una matriz de transición con columnas todas iguales al vector de probabilidad q. Esta matriz tiene la propiedad de que, si x es cualquier vector de probabilidad, se cumple que:

$$Qx = \begin{bmatrix} q_1 & q_1 & \cdots & q_1 \\ q_2 & q_2 & \cdots & q_2 \\ \vdots & \vdots & & \vdots \\ q_k & q_k & \cdots & q_k \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} q_1 + x_1 + q_1 + x_2 + \cdots + q_1 + x_k \\ q_2 + x_1 + q_2 + x_2 + \cdots + q_2 + x_k \\ \vdots \\ q_k + x_1 + q_k + x_2 + \cdots + q_1 + x_k \end{bmatrix}$$

$$(x_1 + x_2 + \cdots + x_k)\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_k \end{bmatrix} = (1)q = q.$$

Es decir, Q transforma a cualquier vector de probabilidad x en un vector de probabilidad fijo q. Entonces, concluye lo siguiente:

Teorema 3.9

Si P es una matriz regular de transición y x es un vector cualquiera de probabilidad, ocurre que cuando $n \to \infty$.

$$P^n x \to \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_k \end{bmatrix} = q$$

Donde q es un vector de probabilidad fijo, independiente de n y con todas las entradas positivas.

Este resultado es, sencillamente, una consecuencia del teorema 3.8 ya que $P^n \to Q$ cuando $n \to \infty$ esto a su vez implica que $P^n x \to Qx = q$ cuando $n \to \infty$.

En consecuencia, en un proceso regular de Markov, el sistema tiende, finalmente, aun vector de estado fijo q. Al vector q se le llama vector de estado permanente o de régimen constante de una cadena regular de Markov.

Para los sistemas con muchos estados, la forma más eficiente de calcular el vector de estado permanente q consiste en hallar para un valor grande de n los tres ejemplos presentados con anterioridad ilustran este procedimiento. Cada uno de ellos es un proceso regular de Markov por lo que es segura la convergencia a un vector de estado permanente. Otra forma de calcular el vector de estado permanente se basa en el siguiente teorema:

Teorema 3.10

El vector de estado permanente q de una matriz regular de transición P, es el vector de probabilidad único que satisface a la ecuación $Pq=q$.

En efecto considérese la identidad matricial $PP^n=P^{n+1}$. Por el teorema 3.7, P^n y P^{n+1} tienden a Q cuando $n \to \infty$ Entonces, $PQ=Q$ y con cualquier columna de Q se obtiene $Pq=q$. Para demostrar que q es el único vector de probabilidad

tal que $Pr=r$. Se debe verificar que $P^n r$ para $n=1, 2, \cdots$. Haciendo que $n \to \infty$ por teorema 3.9 se obtiene $q=r$.

El teorema 3.10 se puede también enunciar diciendo que el sistema lineal homogéneo $(I-P)q=0$ tiene como solución un vector único, q tal que $q_1+q_2+\cdots+q_k=1$. Con base en este nuevo enunciado se calcularán los vectores de estado permanente de los tres ejemplos anteriores.

Ejemplo 3.28 (ya presentado con anterioridad)
En el ejemplo 3.23, la matriz de transición estaba dada por:

$$P=\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix},$$

y por lo tanto el sistema lineal $(I-P)q=0$ es:

$$\begin{bmatrix} .2 \\ -.2 \end{bmatrix}\begin{bmatrix} -.3 \\ .3 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

lo cual conduce a una sola ecuación independiente:

$$.2q_1 - .3q_2 = 0$$

O sea: $q_1 = 1.5q_2$

Haciendo $q_2=s$ cualquier solución de (3.29) tiene la forma $q=s\begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$ donde s es una constante arbitraria. Para que q sea un vector de probabilidad, se hará s=1(1.5+1)=.4 y entonces:

$$q=\begin{bmatrix} .6 \\ .4 \end{bmatrix}$$

Es el vector de estado permanente en este proceso regular de Markov. Lo anterior significa que, a la larga, el 60% de los alumnos hará un donativo en cualquier año, y que el 40% no lo hará. Observe que este resultado es congruente con el obtenido numéricamente cuando se presentó este ejemplo.

Ejemplo 3.29 (ya presentado con anterioridad)
En el ejemplo 3.25, la matriz de transición era:

$$P=\begin{bmatrix} .8 & .3 & .2 \\ .1 & .2 & .6 \\ .1 & .5 & .2 \end{bmatrix},$$

Por lo tanto, el sistema lineal $(I–P)q=0$ es:

$$\begin{bmatrix} .2 & -.3 & -.2 \\ -.1 & .8 & -.6 \\ -.1 & -.5 & .8 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

La forma reducida de la matriz de los coeficientes es (comprobar):

$$\begin{bmatrix} 1 & 0 & -^{34}/_{13} \\ 0 & 1 & -^{14}/_{13} \\ 0 & 0 & 0 \end{bmatrix}$$

Por lo tanto, el sistema lineal origina equivale al sistema:

$$q_1=\left(\frac{34}{13}\right)q_3$$

$$q_2=\left(\frac{14}{13}\right)q_3$$

Haciendo $q_3=s$, cualquier solución de este sistema tiene la forma:

$$q= s\begin{bmatrix} ^{34}/_{13} \\ ^{14}/_{13} \\ 1 \end{bmatrix}$$

Para que s sea un vector de probabilidad, se hará $s=\dfrac{1}{\frac{34}{13}+\frac{14}{13}+1}=\dfrac{13}{61}$ y el vector de estado permanente del sistema será:

$$q=\begin{bmatrix} ^{34}/_{61} \\ ^{14}/_{61} \\ ^{13}/_{61} \end{bmatrix}=\begin{bmatrix} .5573\cdots \\ .2295\cdots \\ .2131\cdots \end{bmatrix}$$

Esto es congruente con el primer resultado, obteniendo numéricamente, del ejemplo 3.29 las entradas de q son a la larga, las probabilidades de que un auto sea devuelto, respectivamente, a las oficinas 1, 2 o 3. Si la agencia que renta automóviles tuviera una flotilla de 1000 unidades, sus instalaciones deberían tener, por lo menos, 558 espacios en la oficina 1,230 en la oficina 2 y 214 en la oficina 3.

Ejemplo 3.30 (ya presentado con anterioridad)
No se presentarán los cálculos en detalle, simplemente se establecerá que el vector único de probabilidad, solución del sistema lineal $(I–P)q=0$ es:

$$q=\begin{bmatrix} 3/28 \\ 3/28 \\ 3/28 \\ 5/28 \\ 4/28 \\ 3/28 \\ 4/28 \\ 3/28 \end{bmatrix} = \begin{bmatrix} .1071\cdots \\ .1071\cdots \\ .1071\cdots \\ .1785\cdots \\ .1428\cdots \\ .1071\cdots \\ .1428\cdots \\ .1071\cdots \end{bmatrix}$$

En este vector, las entradas representan, a la larga, el tiempo que el policía está en cada crucero. Así, si su objetivo es estar el mismo número de horas en cada crucero, la estrategia de cambiase al azar, con iguales probabilidades, de un crucero a otro, no es la correcta.

3.13. NOTAS HISTÓRICAS



Método de solución de Gauss se llama así debido a que Gauss lo describió en un artículo detallando los cálculos que hizo para determinar la órbita del asteroide Pallas. Los parámetros de la órbita tenían que determinarse mediante observaciones del asteroide durante el periodo de seis años comprendido entre 1803 y 1809. Esto dio lugar a seis ecuaciones con seis incógnitas con coeficientes bastante complicados. Gauss mostró cómo

resolver estas ecuaciones reemplazándolas sistemáticamente por un nuevo sistema en el que solo la primera ecuación tenía seis incógnitas, la segunda tenía cinco, la tercera solo cuatro y así sucesivamente, hasta que la sexta ecuación tenía solo una incógnita. Evidentemente, esta última ecuación se podía resolver con facilidad, las incógnitas restantes se hallaron entonces por sustitución regresiva.

La parte Jordán del método de Gauss-Jordán es, en esencia, una técnica sistemática de "sustitución regresiva". En esta forma la describió Wilhelm Jordán (1842-1899), un profesor alemán de geodesia, en su "Manual de geodesia". Este trabajo se publicó por primera vez en alemán en 1873 y desde entonces ha tenido diez ediciones, además de traducciones a otras lenguas. Por otro lado, el propio Jordán atribuye el método a Friedrich Robert Helmert y Peter Andreas Hansen, unos años antes.

Wilhelm Jordán fue importante en su campo a finales del siglo XIX; participó en varias exploraciones geodésicas en Alemania, así como en la primera exploración importante en el desierto de Libia. También fue editor fundador de la revista alemana de geodesia. Su interés por hallar un método sistemático para resolver grandes sistemas de ecuaciones surge de la frecuente aparición de éstos en problemas de triangulación.

Un método de reducción matricial para resolver un sistema de ecuaciones lineales se encuentra en el antiguo trabajo chino "Nueve capítulos del arte matemático". El autor plantea la siguiente solución al sistema.

$$3x + 2y + z = 39$$
$$2x + 3y + z = 34$$
$$x + 2y + 3z = 26$$

El diagrama de los coeficientes se dispone en un ábaco como sigue:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 26 & 34 & 39 \end{bmatrix}$$

Lo que nos induce a multiplicar la columna del centro por 3 y después restar a la columna de la derecha "el máximo de veces posible"; lo mismo se hace con la

columna de la izquierda. Los nuevos diagramas son entonces:

$$\begin{bmatrix} 1 & 0 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 1 \\ 26 & 24 & 39 \end{bmatrix} \approx \begin{bmatrix} 0 & 0 & 3 \\ 4 & 5 & 2 \\ 8 & 1 & 1 \\ 39 & 24 & 39 \end{bmatrix}$$

La siguiente instrucción es multiplicar la columna de la izquierda por 5 y restar de nuevo la columna del centro el máximo de veces posible, lo que da:

$$\begin{bmatrix} 0 & 0 & 3 \\ 0 & 5 & 2 \\ 36 & 1 & 1 \\ 99 & 24 & 39 \end{bmatrix}$$

Así, el sistema se redujo al sistema $3x+2y+z=39$, $5y+z=24$, $36z=99$, del cual se halla fácilmente la solución completa.

La noción de la inversa de una matriz apareció por primera vez en 1855 en un escrito de Arthur Cayley (18211895) y se amplió en un artículo publicado tres años más tarde titulado “Una memoria sobre la teoría de matrices”. En este trabajo, Cayley describe las propiedades básicas de las matrices, puntualizando que la mayoría se deriva del trabajo con conjuntos de ecuaciones lineales. En particular, la inversa procede de la idea de resolver un sistema:

$$X = ax + by + cz$$
$$Y = a'x + b'y + c'z$$
$$Z = a''x + b''y + c''z$$

Para x, y, z en función de X, Y, Z. Cayley da una construcción explícita para la inversa en función de los determinantes de la matriz original y de los menores.

En 1842, **Arthur Cayley** se graduó en el Trinity College, Cambridge, pero



no pudo hallar un puesto adecuado de profesor por lo que, al igual que Sylvester, estudió leyes y se hizo abogado en 1849. Durante sus catorce años de ejercicio escribió alrededor de 300 artículos matemáticos; finalmente, en 1863 se hizo profesor de Cambridge, donde permaneció hasta su muerte. Fue durante su época de abogado cuando

conoció a Sylvester; sus discusiones durante los 40 años siguientes fueron muy fructíferas para el progreso del álgebra. Durante su vida, Cayley escribió alrededor de 1000 artículos sobre matemática pura, dinámica teórica y astronomía matemática.

3.14. EJERCICIOS PROPUESTOS

1. Demuestre que la eliminación de Gauss-Jordan tiene una magnitud de operaciones de orden $n^{3/2}$.
2. Demuestre que el proceso para encontrar la inversa de una matriz tiene una magnitud de operaciones de orden n^3.
3. Para una matriz de $m \times n$ diga qué condiciones se deben cumplir para que haya solución única, soluciones múltiples o no haya solución.
4. Determine la factorización LU de la siguiente matriz:

$$A = \begin{bmatrix} 6 & 2 & 1 & -1 \\ 2 & 4 & 1 & 0 \\ 1 & 1 & 4 & -1 \\ -1 & 0 & -1 & 3 \end{bmatrix}$$

5. Suponga que se ha encontrado la factorización $A=LU$, pero el sistema por resolver es $A^T y=b$ en lugar de $Ax=b$. Encuentre una manera rápida de calcular y.
6. Encuentre la inversa de las siguientes matrices, si existe:

a) $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

b) $B = \begin{bmatrix} 1 & 1 & \cdots & \cdots & 1 & 1 \\ 0 & 1 & \cdots & \cdots & 1 & 1 \\ 0 & 0 & 1 & \cdots & 1 & 1 \\ 0 & 0 & \cdots & \cdots & 1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

7. En un artículo reciente sobre zoología, se reportó que la longitud promedio de ala que resulta de aparear tres variedades de mutantes de las moscas de fruta (Drosophila melanogaster) puede expresarse en la forma de la matriz simétrica:

$$A=\begin{bmatrix}1.59 & 1.69 & 2.13\\ 1.69 & 1.31 & 1.72\\ 2.13 & 1.72 & 1.85\end{bmatrix}$$

Donde a_{ij} denota la longitud promedio de la cría resultante de aparear un macho del tipo i con una hembra del tipo j.

a) ¿Qué significado físico se asocia con la simetría de la matriz?
b) ¿Es esta matriz positiva definida? Si es así, demuéstrelo; si no, encuentre un vector x diferente de cero para el cual $x^tAx\leq 0$.

8. Resuelva el siguiente sistema usando el método de Gauss con sustitución regresiva.

$x_1+2x_2-x_3= -3$
$3x_1+7x_2+2x_3= 1$
$4x_1-2x_2+x_3= -2$

Encuentre la solución única del siguiente sistema usando el método de Gauss-Jordan.

$x_1+2x_2 - x_3= 8$
$2x_1+5x_2-6x_3= 17$
$-x_1-2x_2+x_3-x_4= -8$
$4x_1+10x_2-9x_3+x_4= 33$

9. **Distribución de población**
Consideremos situaciones en las que se separa a la gente en dos o más categorías. Por ejemplo, podemos separar a los ciudadanos de un país, de acuerdo con sus ingresos, en categorías de pobre, ingresos medios, rico.
Podemos separar a los habitantes de la República Mexicana en categorías

de acuerdo con el clima en el que viven: caluroso, templado, frío.
Como terminología, hablaremos de una población separada en estados. Para los dos ejemplos anteriores, la población y los estados están dados por lo siguiente:

Población	Estados
Ciudadanos de un país	pobre, ingreso medio, rico
Habitantes de la Rep. Mex.	caluroso, templado, frío

Estamos interesados en cómo la distribución de una población entre estados puede cambiar durante un periodo de tiempo. Las matrices y su multiplicación pueden desempeñar un papel importante en dichas consideraciones.

La tendencia de una población a moverse entre n estados se puede describir a veces mediante una matriz de $n \times n$.

Consideremos una población distribuida entre n=3 estados, que llamaremos estado 1, estado 2 y estado 3. Se supone que conocemos la proporción de la población del estado j, que se mueve al estado i en un determinado periodo de tiempo fijo. Observe que la dirección del movimiento del estado j al estado i va en orden de derecha a izquierda de los subíndices en . La matriz se llama matriz de transición.

Ejemplo. Suponga que la población de un país está clasificada de acuerdo con los ingresos en:
Estado 1: pobre
Estado 2: ingresos medios
Estado 3: rico

Y suponga también que en cada periodo de 20 años (alrededor de una generación) tenemos los siguientes datos para la población y su descendencia:

De la gente pobre, el 19% pasó a ingresos medios, y el 1% a rica; de la gente con ingresos medios, el 15% pasó a pobre, y el 10% a rica; de la gente rica, el 5% pasó a pobre y el 30% a ingresos medios.

La matriz de transición que describe estos datos se construye de la siguiente manera:

El registro de la matriz de transición *T* representa la proporción de la población que pasa del estado *j* al estado *i*, no el porcentaje. Así el dato de que el 19% de los pobres (estado 1) pasará a ingresos medios (estado 2) significa que debemos tomar t_2= 0.19. Igualmente, como el 1% de la gente del estado 1 pasa al estado 3 (rico), debemos tomar t_{31}= 0.01 y así sucesivamente obteniéndose la matriz:

$$
\begin{array}{c c l}
 & \begin{array}{ccc} \textit{pobre} & \textit{medio} & \textit{rico} \end{array} & \\
T= & \begin{bmatrix} 0.80 & 0.15 & 0.05 \\ 0.19 & 0.75 & 0.30 \\ 0.01 & 0.10 & 0.65 \end{bmatrix} & \begin{array}{l} \textit{pobre} \\ \textit{medio} \\ \textit{rico} \end{array}
\end{array}
$$

Hemos dado a las columnas y los renglones los nombres de los estados. Obsérvese que un registro de la matriz da la proporción de la población del estado superior del registro que pasa al estado de la derecha del registro durante un período de 20 años. La suma de registros de cada columna de *T* es 1, pues la suma refleja el movimiento de toda la población para el estado relacionado en la parte superior de la columna.

Supongamos en este ejemplo, que las proporciones de toda la población ubicada en los distintos estados al comienzo de un periodo de tiempo, están dadas en el vector columna:

$$
P= \begin{bmatrix} p1 \\ p2 \\ p3 \end{bmatrix}
$$

Por ejemplo, se podría tener $p=[\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3}]_t$ si el total de la población estuviera inicialmente dividida por igual entre los estados. Los registros en ese vector *p* de distribución de la población deben ser no negativos y sumar 1.

Hallemos la proporción de toda la población que se encuentra en el estado 1 después de un periodo de tiempo de 20 años, a sabiendas de que inicialmente la proporción en el estado 1 es p_1. La proporción de la población del estado 1 que permanece en el estado 1 es t_{11}. Esto da una contribución de $t_{11}p_1$ a la proporción de toda la población que se encontrará en el estado 1 al cabo de 20 años. Lógicamente también tendremos contribuciones de los estados 2 y 3 al estado 1 al cabo de 20 años. Así después de 20 años, la

proporción en el estado 1 es:
$t_{11} p_1 + t_{12} p_2 + t_{13} p_3$

Observe que este es precisamente el primer registro en el vector columna dado por el producto:

$$Tp = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

De manera similar se encuentra que la segunda y tercera componentes de Tp dan las proporciones de población del estado 2 y del estado 3 después de un periodo de tiempo.

En el ejemplo se halló una matriz de transición que gobierna el flujo de una población entre tres estados en un periodo de 20 años. Supongamos que la misma matriz de transición es válida para el siguiente periodo de 20 años, y para los siguientes veinte años y así sucesivamente, esto es, existe una sucesión o cadena de periodos de 20 años para los cuales es válida la matriz de transición.

Una situación así se denomina CADENA DE MARKOV. Y se define una matriz de transición de m periodos como: Una cadena de Markov con matriz de transición T tiene T^m como matriz de transición para m periodos.

Clasifiquemos a las mujeres de un país según vivan en un área urbana (U), suburbana(S) o rural (R). Se supone, además, que:

Para las mujeres urbanas, el 10% de sus hijas se establecen en áreas rurales, y el 50% en áreas suburbanas, para las mujeres suburbanas, el 20% de sus hijas se establecen en áreas rurales, y el 30% en áreas urbanas; para las mujeres rurales el 20% se establece en los suburbios, y el 70% en áreas rurales. Sea el periodo de esta cadena de Markov, el tiempo para producir a la siguiente generación.

a) Encuentre la matriz de transición para esta cadena de Markov, tomando los estados en el orden *U, S, R*.
b) Halle la proporción de mujeres rurales cuyas nietas son también rurales.
c) Si el vector de distribución de población inicial para todas las mujeres es hallar el vector de distribución de población para la generación siguiente.

10. **Modelos económicos de Leontief**

Tres vecinos tienen hortalizas en sus jardines. El vecino *A* cultiva jitomates, *B* cultiva maíz y *C* lechugas. Los tres se ponen de acuerdo para repartir la cosecha entre ellos de la siguiente manera: *A* se quedará con la mitad de sus jitomates y recibirá ⅓ de la cosecha de maíz y ¼ de la de lechuga. *B* recibirá ⅓ de la cosecha de jitomate, ⅓ de la de maíz y ¼ de la de lechuga, *C* un ⅙ de la de jitomate, ⅓ de la de maíz y se quedará con la mitad de su cosecha de lechuga ¿Qué precios deberán asignar a los respectivos productos si se debe satisfacer la condición de equilibrio de una economía cerrada y si la cosecha de menor precio debe venderse a $1000?

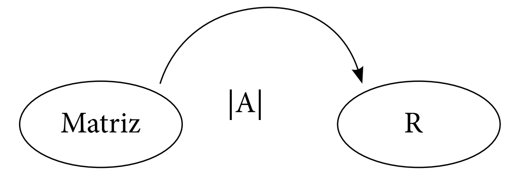
CONTENIDO

Capítulo 4
DETERMINANTES

4.1. INTRODUCCIÓN

En el capítulo dos hemos visto cómo podemos representar en forma muy sencilla sistemas de ecuaciones lineales a través de matrices. Asociadas a dichas matrices, existe otro concepto que es el que nos ocupa en este capítulo y es el de determinantes. Sus usos y aplicaciones son variados, como podrá verse más adelante. Podemos adelantar sin embargo que el objetivo de los determinantes ha sido el de simplificar cálculos numéricos asociados a los sistemas de ecuaciones lineales, sobre todo cuando son muy grandes.

Definición 4.1
Un determinante es una función que asigna un número a una matriz cuadrada.



Comenzaremos este tema con una introducción a determinantes de 2×2 y de 3×3 motivada por cálculos de área y volumen.

4.2. ÁREA DE UN PARALELOGRAMO

En la figura siguiente se muestra el paralelogramo determinado por dos vectores distintos de cero, no paralelos, a=(a_1, a_2) y b=(b_1, b_2) de R^2. Este paralelogramo tiene un vértice en el origen; consideramos las flechas que representan *a* y *b* como los dos lados del paralelogramo, con el origen como vértice común.

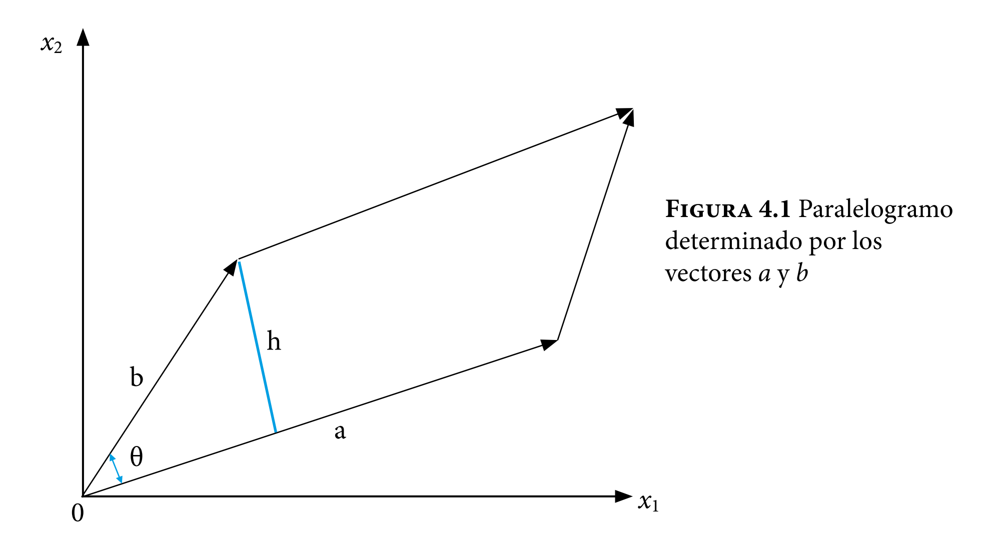


Figura 4.1 Paralelogramo determinado por los vectores *a* y *b*

Podemos hallar el área de este paralelogramo multiplicando la longitud $\|a\|$ de su base por su altura h obteniendo:

$$\text{Área } (\square) = \|a\|\, h = \|a\| \; \|b\| \,(\text{sen}\Theta) = \|a\| \; \|b\| \sqrt{}[1 - cos^2\Theta] \tag{4.1}$$

Observación:
Recuerde que la magnitud de un vector a en R^n se define:

$$\|a\| = \sqrt{}[a_1{}^2 + a^2{}_2 + \cdots + a_n{}^2]$$

Y el ángulo entre dos vectores *a* y *b* como:

$$Arc\ cos = \left[\frac{a \qquad b}{\|a\| \quad \|b\|} \right]$$

Donde se aplicó la ley de los cosenos:

$$\|a\| \, \|b\| \, cos\theta = a_1 b_1 + \cdots + a_n b_n$$

Ejemplo 4.1
Encuentre el ángulo Θ entre los vectores (1, 2, 0, 2) y (−3, 1, 1, 5) en R^4 .

Solución
Tenemos:

$$\cos\theta = \frac{(1,2,0,2)\cdot(-3,1,1,5)}{\sqrt{1^2+2^2+0^2+2^2}\,\sqrt{(-3)^2+1^2+5^2}} = \frac{9}{18} = \frac{1}{2}$$

Así cosθ = 60º.

Volviendo a la ecuación (4.1) tenemos que si usamos el resultado de que: $ab = ||a||\ ||b|| \cos\theta$ y elevamos al cuadrado la ecuación:

$$\begin{aligned}(\text{Área})^2 &= \|a\|^2\ \|b\|^2 - \|a\|^2\ \|b\|^2 \cos^2\theta = \|a\|^2\ \|b\|^2 - (a\ b)^2 \\ &= (a_1{}^2 + a_2{}^2)\ (b_1{}^2 + b_2{}^2) - (a_1\ b_1 + a_2 b_2) \\ &= (a_1\ b_2 - a_2\ b_1)^2\end{aligned}$$

Tomando raíces cuadradas tenemos:

Área (▱) $= |a_1\ b_2 - a_2\ b_1| = det$ A

Donde A es la matriz:

$$A = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$$

Ejemplo 4.2
Encuentre el área del paralelogramo en R^2 con vértices (1, 1), (2, 3), (2, 1), (3, 3).

Solución
El paralelogramo se ilustra en la figura siguiente. Los lados que tienen (1,1) como vértice común se pueden considerar como los vectores a y b:

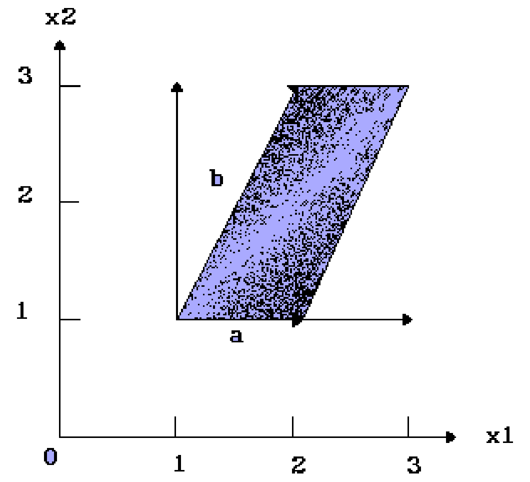


FIGURA 4.2 Paralelogramo determinado por los vectores *a* y *b*

De aquí podemos encontrar que *a* y *b* están dados por:

$a= (2,1) - (1,1) = (1,0)$
$b= (2,3) - (1,1) = (1,2)$

Por lo tanto, el área del paralelogramo está dada por el determinante:

$$\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = (1)(2) - (0)(1) = 2$$

4.3. PRODUCTO CRUZ

El determinante asociado con una matriz cuadrada de 2×2 se denomina **determinante de segundo orden**, otra aplicación de estos determinantes aparece cuando buscamos un vector de R^3 que sea perpendicular a cada uno de los dos vectores independientes dados:

$b= (b_1, b_2, b_3)$ y $c = (c_1, c_2, c_3)$

Los vectores coordenados unitarios de R^3 son:

$i= (1,0,0)$ $j= (0,1,0)$ y $k= (0, 0, 1)$

Entonces se verifica que:

$$p= \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} i - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} j + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} k \qquad (4.2)$$

Es un vector ortogonal a b y c. Esto se puede observar calculando $pb = pc = 0$.

El vector p se conoce como producto cruz de b y c, y se denota por $p = b{\times}c$. Hay una forma sencilla de recordar la fórmula (4.2) para el producto cruz $b{\times}c$. Se forma la siguiente matriz de 3×3:

$$\begin{bmatrix} i & j & k \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

Cuyo determinante es igual a:

$$\begin{bmatrix} i & j & k \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} = i \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - j \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + k \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

Que no es otra que la fórmula (4.2).

Ejemplo 4.3

Encuentre un vector perpendicular a (2, 1, 1) y (1, 2, 3) de R^3.

Solución

Se forma la matriz: $\begin{bmatrix} i & j & k \\ 2 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$

Y se encuentra que:

$$(2,1,1)\ x\ (1,2,3) = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} i - \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} j + \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} k$$

$$= i - 5j + 3k = (1,0,0) - 5\ (0,1,0) + 3\ (0,0,1)$$
$$= (1, -5, 3)$$

El producto cruz $p = b \times c$ no solo es perpendicular a b y c, sino que apunta en la dirección determinada por la conocida **regla de la mano derecha:** cuando los dedos de la mano derecha se curvan en la dirección de b a c, entonces el pulgar apunta en la dirección de $b \times c$ como en la siguiente figura:

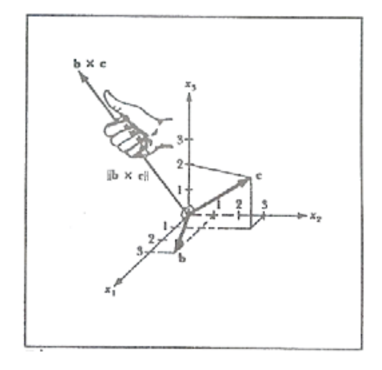


Figura 4.3 Ilustración de la regla de la mano derecha. Fuente Fraleigh (1989)

También tiene interés la magnitud del vector $p = b \times c$ de (4.2): es el área del paralelogramo con un vértice en el origen de R^3 y aristas en ese vértice dadas por los vectores b y c. Para ver esto se considera un diagrama como el de la figura 4.1, pero con b reemplazado por c y se calcula nuevamente el área como sigue:

$$\begin{aligned}(\acute{A}rea)^2 &= \|c\|^2 \ \ \|b\|^2 - (c\ b)^2 \\ &= (c_1^2 + c_2^2 + c_3^2)\ (b_1^2 + b_2^2 + b_3^2) - (c_1\ b_1 + c_2\ b_2 + c_3\ b_3)^2 \\ &= \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} + \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}^2\end{aligned}$$

Tomando raíces cuadradas obtenemos:

El área del paralelograma en R^3 dado por b y $c = \| b \times c \|$.

Ejemplo 4.4

Encuentre el área del paralelogramo en R^3determinado por los vectores b=(3, 1, 0) y c=(1, 3, 2).

Solución

Se forma la matriz: $\begin{bmatrix} i & j & k \\ 3 & 1 & 0 \\ 1 & 3 & 2 \end{bmatrix}$

De aquí:

$$b \times c = \begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix} i - \begin{vmatrix} 3 & 0 \\ 1 & 2 \end{vmatrix} j + \begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} k = (2,-6,8)$$

Por lo tanto: $\| b \ x \ c \| = 2\sqrt{\left[(1)^2 + (-3)^2 + (4)^2\right]} = 2\sqrt{26}$

Ejemplo 4.5

Encuentre el área del triángulo en R^3 con vértices (-1, 2, 0), (2, 1, 3) y (1, 1, -1)

Solución

Si consideramos (−1, 2, 0) como un origen local y tomamos los vectores correspondientes a flechas que comiencen ahí y lleguen a (2, 1, 3) y (1, 1, −1), a saber:

$a = (2, 1, 3) - (-1, 2, 0) = (3, -1, 3)$
$b = (1, 1, 1) - (-1, 2, 0) = (2, -1, -1)$

Ahora $\|a \times b\|$ es el área del paralelogramo determinado por estos vectores y el área del triángulo es la mitad del área del paralelogramo:

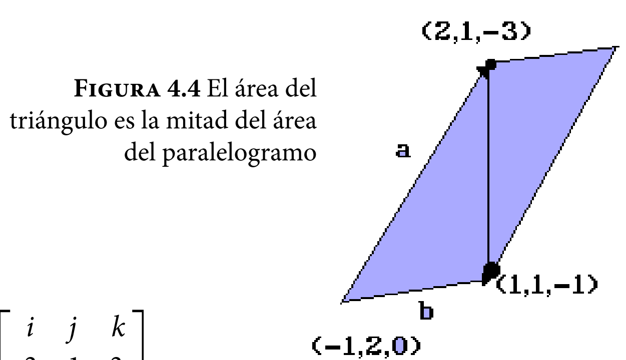


Figura 4.4 El área del triángulo es la mitad del área del paralelogramo

Se forma la matriz: $\begin{bmatrix} i & j & k \\ 3 & -1 & 3 \\ 2 & -1 & -1 \end{bmatrix}$

De aquí:

$$a \times b = \begin{vmatrix} -1 & 3 \\ -1 & -1 \end{vmatrix} i - \begin{vmatrix} 3 & 3 \\ 2 & 1 \end{vmatrix} j + \begin{vmatrix} 3 & -1 \\ 2 & -1 \end{vmatrix} k = (4, 9, -1)$$

Por lo tanto: $\|a \times b\| = \sqrt{\left[(4)^2 + (9)^2 + (-1)^2\right]} = \sqrt{98} = 7\sqrt{2}$ de aquí que el área del triángulo es $7\sqrt{2/2}$.

4.4. VOLUMEN DE UNA CAJA

El producto cruz es útil para hallar el volumen de la caja o paralelepípedo determinado por tres vectores distintos de cero $a= (a_1, a_2, a_3)$, $b= (b_1, b_2, b_3)$ y $c= (c_1, c_2, c_3)$ en R^3 como se muestra en la figura 4.5.

El volumen de la caja se puede calcular multiplicando el área de la base por la altura h: el área de la base de la caja es igual a $\|b\times c\|$ y la altura se puede encontrar calculando

$$h= \|a\|\ |cos\theta| = \frac{\|b\times c\|\ \|a\|\ |cos\theta|}{\|b\times c\|} = \frac{|(b\times c)a|}{\|b\times c\|}$$

El valor absoluto se usa en el caso en que cos Θ es negativo. Este sería el caso si *bxc* tuviera dirección opuesta a la que se muestra en la figura 4.5. Así:

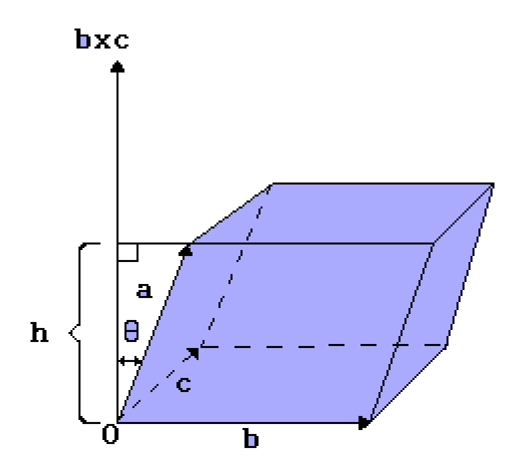


Figura 4.5 La caja en R^3 determinada por *a*, *b* y *c*

Volumen (caja) = área de la base × altura

$$= \frac{\|b\times c\|\ |(bcx)a|}{b\times c}$$

$$= |(b\times c)a|$$

Es decir; el volumen queda definido como:

$$\text{Volumen} = |a_1(b_2c_3-b_3c_2) - a_2(b_1c_3-b_3c_1) + a_3(b_1c_2-b_2c_1)|$$
$$= \det A$$

Donde:

$$A=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

Ejemplo 4.6

Encuentre el volumen de la caja con vértice en el origen, determinada por los vectores a= (4, 1, 1), b= (2, 1, 0) y c= (0, 2, 3).

Solución

La caja se muestra en la figura 4.6, y se forma la matriz, cuyos renglones son los vectores a, b y c. Su volumen está dado por el valor absoluto del determinante:

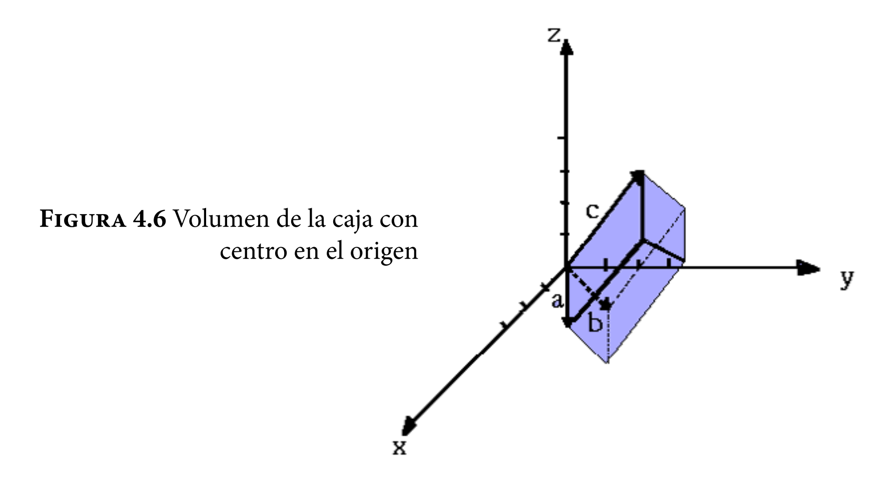


Figura 4.6 Volumen de la caja con centro en el origen

$$\begin{vmatrix} 4 & 1 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 3 \end{vmatrix} = 4\begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} - \begin{vmatrix} 2 & 0 \\ 0 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix} = 10$$

4.5. EL DETERMINANTE DE UNA MATRIZ CUADRADA

El determinante de una matriz de 1×1 es su único registro. Así, un determinante de segundo orden se puede definir como ya se dijo:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\, a_{22} - a_{21}\, a_{a12}$$

Y un determinante de tercer orden en términos de uno de segundo orden, como:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

O bien:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}\, a_{21}\, a_{33}$$

$- \quad - \quad - \qquad + \quad + \quad +$

Definimos un determinante de n-ésimo orden en términos de determinantes de orden n–1. Para lo cual se introduce la matriz menor A_{ij} de una matriz A de n×n, que es la matriz de (n–1)×(n–1) obtenida al eliminar la i-ésima fila y la j-ésima columna de A.

$$Aij = \begin{bmatrix} a_{11}\cdots & a_{1j}\cdots & a_{1n} \\ a_{i1}\cdots & a_{i1}\cdots & a_{in} \\ a_{n1}\cdots & a_{nj}\cdots & a_{nn} \end{bmatrix} \Rightarrow i\text{– ésima fila} \qquad (4.3)$$

$\Downarrow$ j-ésima columna

Usando $|A_{ij}|$ como notación para el determinante de la matriz menor A_{ij} podemos expresar el determinante de una matriz A de 3×3 como:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\,|A_{11}| - a_{12}\,|A_{12}| + a_{13}\,|A_{13}|$$

Los números $a_{11'} = |A_{11}|$, $a_{12'} = -|A_{12}|$ y $a_{13'} = |A_{13}|$ son apropiadamente llamados **cofactores** de a_{11}, a_{12} y a_{13}.

Definición 4.2

El determinante de una matriz de 1×1 es su único registro. Sea $n>1$ y supongamos definidos los determinantes de orden menor que n. Sea $A=(A_{ij})$ una matriz de orden $n \times n$. El cofactor de a_{ij} en A es:

$$a_{ij}' = (-1)^{i+j} \, det \, (A_{ij})$$

Donde A_{ij} es la matriz menor de A dada en (4.3). El determinante de A es:

$$\det A = \begin{vmatrix} a_{11} \cdots \cdots\cdots \cdots a_{1n} \\ a_{il} \cdots \cdots\cdots \cdots a_{in} \\ a_{nl} \cdots \cdots\cdots \cdots a_{nn} \end{vmatrix} = a_{11}\, a_{11'} + a_{12}\, a_{12'} + \cdots\cdots + a_{1n}\, a_{1n'}$$

Ejemplo 4.7

Encuentre el cofactor del registro 3 o entrada a_{21} en la matriz:

$$A = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 2 \\ 4 & 0 & 1 & 4 \\ 1 & 0 & 2 & 1 \end{bmatrix}$$

Solución

$$a'_{21} = (-1)^{2+1} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 2 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 4 \\ 2 & 1 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 0 & 2 \end{vmatrix} = 7$$

Ejemplo 4.8

$$A = \begin{bmatrix} 5 & -2 & 4 & -1 \\ 0 & 1 & 5 & 2 \\ 1 & 2 & 0 & 1 \\ -3 & 1 & -1 & 1 \end{bmatrix}$$

Usando la definición 4.3, encuentre el determinante de la matriz:

Solución

$$det\ A=\begin{bmatrix} 5 & -2 & 4 & -1 \\ 0 & 1 & 5 & 2 \\ 1 & 2 & 0 & 1 \\ -3 & 1 & -1 & 1 \end{bmatrix}$$

$$= 5(-1)^2\begin{vmatrix} 1 & 5 & 2 \\ 2 & 0 & 1 \\ 1 & -1 & 1 \end{vmatrix} +(-2)(-1)^3\begin{vmatrix} 0 & 5 & 2 \\ 1 & 0 & 1 \\ -3 & -1 & 1 \end{vmatrix} +4(-1)^4\begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ -3 & 1 & 1 \end{vmatrix} +(-1)(-1)^5\begin{vmatrix} 0 & 1 & 5 \\ 1 & 2 & 0 \\ -3 & 1 & -1 \end{vmatrix}$$

Calculando los determinantes de tercer orden se tiene:

$$\begin{vmatrix} 1 & 5 & 2 \\ 2 & 0 & 1 \\ 1 & -1 & 1 \end{vmatrix} = 1\begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} -5\begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} +2\begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} = 8$$

$$\begin{vmatrix} 0 & 5 & 2 \\ 1 & 0 & 1 \\ -3 & -1 & 1 \end{vmatrix} = 0\begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} -5\begin{vmatrix} 2 & 1 \\ -3 & 1 \end{vmatrix} +2\begin{vmatrix} 1 & 0 \\ -3 & -1 \end{vmatrix} = -22$$

$$\begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ -3 & 1 & 1 \end{vmatrix} = 0\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} -1\begin{vmatrix} 1 & 1 \\ -3 & 1 \end{vmatrix} +2\begin{vmatrix} 1 & 2 \\ -3 & -1 \end{vmatrix} = 10$$

$$\begin{vmatrix} 0 & 1 & 5 \\ 1 & 2 & 0 \\ -3 & 1 & -1 \end{vmatrix} = 0\begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} -1\begin{vmatrix} 1 & 0 \\ -3 & -1 \end{vmatrix} +5\begin{vmatrix} 1 & 2 \\ -3 & 1 \end{vmatrix} = 36$$

De donde det $(A) = 5(-8) + 2(-22) + 4(10) + 1(36) = -8$

Teorema 4.1: Desarrollo general por menores

Sean A una matriz de 2×2 y r y s cualesquiera selecciones de la lista de números $1, 2,\cdots,n$ entonces

$$\det(A) = a_{r1}\,a'_{r1} + a_{r2}\,a'_{r2} + \cdots + a_{rn}\,a'_{rn} \quad (4.4)$$

$$= a_{1s}\,a'_{1s} + a_{2s}\,a'_{2s} + \cdots + a_{ns}\,a'_{ns} \quad (4.5)$$

La ecuación (4.4) es el desarrollo del determinante de A por menores en la r-ésima fila de A y la ecuación (4.5) es el desarrollo del determinante de A por menores en la s-ésima columna de A. Este teorema establece que se puede hallar el determinante de A desarrollando por menores en cualquier fila o columna de A.

Ejemplo 4.9

Encuentre el determinante de la matriz:

$$A=\begin{bmatrix} 3 & 2 & 0 & 1 & 3 \\ -2 & 4 & 1 & 2 & 1 \\ 0 & -1 & 0 & 1 & -5 \\ -1 & 2 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

Solución

Se puede agilizar el cálculo desarrollando por menores en cada paso, en la fila o columna que contenga la mayoría de ceros:

$$det\,(A)= 2(1)^{5+5}\begin{vmatrix} 3 & 2 & 0 & 1 \\ -2 & 4 & 1 & 2 \\ 0 & -1 & 0 & 1 \\ -1 & 2 & 0 & -1 \end{vmatrix} \qquad \text{Desarrollando en la 5ª fila}$$

$$= 2(1)(-1)^{2+3}\begin{vmatrix} 3 & 2 & 1 \\ 0 & -1 & 1 \\ -1 & 2 & -1 \end{vmatrix} \qquad \text{Desarrollando en la 3ª columna}$$

$$= -2\left[3(-1)^{1+1}\begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} -1(-1)^{3+1}\begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix}\right] \qquad \text{Desarrollando en la 1ª columna}$$

$$=12$$

El uso de esta definición para determinantes de 4×4 ya es un trabajo tremendo, pues se involucran 4! productos, como se podrá observar en la sección 4.7, por lo cual se usan propiedades de los determinantes para calcularlos más fácilmente, esto se verá en la siguiente sección.

PROPIEDADES DE LOS DETERMINANTES

1. Si cualquier renglón o columna de A, tiene solo componentes cero entonces:

   det $(A) = 0$.

   Por ejemplo: $A=\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ Entonces, $det\ A= a(0) - b(0) = 0$

2. Si A tiene dos renglones (o columnas) iguales o proporcionales, entonces det $(A)=0$

   Por ejemplo: $A=\begin{bmatrix} a & b \\ 2a & 2b \end{bmatrix}$ Entonces det $A= 2ab - 2ab = 0$

3. Sea B una matriz que resulta de intercambiar dos renglones (columnas) de la matriz cuadrada A, entonces:

   det $B = -$det A.

   Por ejemplo:

   $A= \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $B= \begin{bmatrix} c & d \\ a & b \end{bmatrix}$

   → det $A= ad-bc$ det $B= cb-ad = -(ad -bc)$

4. Si A es una matriz cuadrada entonces:

   det A = det A^t

5. Si B es la matriz que resulta de multiplicar por α un renglón i de A entonces:

   $det\ B = \alpha$ det A

Por ejemplo:

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B=\begin{bmatrix} \alpha a & \alpha b \\ c & d \end{bmatrix}$$

$\rightarrow$ *det A = ad–bc* $\qquad$ *det B* $= \alpha ad - \alpha bc$

$= \alpha(ad - bc)$

6. Sea $B = \alpha A$, donde A es una matriz de $n\times n$ entonces det $B = \alpha^n$ det A.

   Por ejemplo:

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B=\begin{bmatrix} \alpha a & \alpha d \\ \alpha c & \alpha d \end{bmatrix}$$

$\rightarrow$ *det A = ad–bc* $\qquad$ *det B* $= (\alpha a)(\alpha d) - (\alpha b)(\alpha c)$

$= \alpha^2(ad) - \alpha^2(bc)$

$= \alpha^2(ad - bc)$

   Como A es de 2×2 entonces $\alpha^n = \alpha^2$.

7. El determinante de la matriz identidad es igual a 1.

8. Si B se obtiene de A por medio de la operación elemental de sumar a un renglón un múltiplo de otro: $E_{pq}^{(\alpha)}$,entonces:

   det B = det A

   Por ejemplo:

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B=\begin{bmatrix} a+\alpha c & b+\alpha d \\ c & d \end{bmatrix}$$

det A= ad – bc $\qquad$ *det B* $= (a+\alpha c)d - (b+\alpha d)c$

$= ad + \alpha cd - bc - \alpha dc$

$= ad - bc$

9. Para A y B de orden $n \times n$ el determinante del producto es igual al producto de los determinantes, entonces:

$$\det(AB) = (\det A)(\det B)$$

Por ejemplo:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

$det\ A = ad - bc$ $\qquad det\ B = eh - gf$

$det\ A\ det\ B = (ad-bc)(eh-gf) = adeh - adgf - bceh + bcgf$

$$AB = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

$det\ AB = (ae+bg)(cf+dh)-(ce+dg)(af+bh) = aedh - afdg + bgcf - bhce$

En particular, si A es invertible, entonces:

$det\ A \det A^{-1} = \det AA^{-1} = \det I = 1$ luego $\det A^{-1} = 1/\det A$

10. Si A es singular, entonces:

$\det A = 0$. Si A es no singular $\det A^{1} \neq 0$.

Ejemplo 4.10

Sea la matriz:

$$A = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}$$

Cuyo determinante es igual a 10. Usando propiedades del determinante calcule:

a) $\det(\alpha A)$
b) $\det(\alpha A^{-1})$
c) $\left[\det(\alpha A)\right]^{-1}$
d) $\det(4AA^{t})$

Solución

e) α^4 (det A) = α^4 (10)

f) α^4 (det A^{-1}) = α^4 (1/det A) = $\alpha^4/10$

g) $(\alpha^4\ 10)^{-1}$

h) 4^4 (10)(10) = 4^4(100)

4.6. CÁLCULO DE DETERMINANTES MEDIANTE LA REDUCCIÓN A UNA MATRIZ TRIANGULAR SUPERIOR O INFERIOR

Hemos visto que el cálculo de determinantes de orden superior es una tarea irracional si se realiza directamente a partir de la definición 4.3, usando solamente desarrollos repetidos por menores; sabemos también que una matriz se puede reducir a una matriz triangular superior al efectuar operaciones elementales por filas. El siguiente teorema nos permite usar esto para calcular determinantes de matrices de orden mayor que 3.

Teorema 4.2

Sea $A = (A_{ij})$ una matriz de orden $n \times n$ triangular superior o inferior, entonces:

$$\det A = a_{11}\, a_{22} \cdots a_{nn}$$

Lo que significa que el determinante de una matriz triangular es igual al producto de sus componentes diagonales.

Demostración

Sea U una matriz triangular superior tal que:

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

Si desarrollamos cada vez en las primeras columnas, tenemos:

$$\det U = u_{11} \begin{vmatrix} u_{22} & \cdots & u_{2n} \\ 0 & u_{33} \cdots & u_{3n} \\ 0 & 0 \cdots & u_{nn} \end{vmatrix} = u_{11}\, u_{22} \begin{vmatrix} u_{33} & \cdots & u_{3n} \\ 0 & u_{44} \cdots & u_{4n} \\ 0 & 0 \cdots & u_{nn} \end{vmatrix} = u_{11}\, u_{22}, \cdots \cdots, u_{nn}$$

CONTENIDO

Además, el efecto de las matrices elementales en los determinantes lo podemos encontrar en la siguiente proposición:

PROPOSICIÓN 4.1

Sea A una matriz de $n \times n$ y E una matriz elemental. Entonces:

$$\det EA = \det E \det A, \text{ o bien } \det AE = \det A \det E.$$

Demostración

Existen tres tipos de matrices elementales:

a) $E = E_{pq}$ sea $B = E_{pq} A$ entonces:

$$\det E_{pq} A = \det B = -\det A = (-1) \det A = \det E_{pq} \det A$$

b) $E = E_p^{(\alpha)}$ sea $B = E_p^{(\alpha)} A$ entonces:

$$\det E_p^{(\alpha)} A = \det B = \alpha \det A = \det E_p^{(\alpha)} \det A$$

c) $E = E_{pq}^{(\beta)}$ se deja de ejercicio.

Ejemplo 4.11

El determinante de la matriz A:

$$A = \begin{bmatrix} 2 & 1 & 7 \\ 0 & 2 & -5 \\ 0 & 0 & 1 \end{bmatrix} \text{ Es igual a } |A| = (2)(2)(1) = 4$$

Ejemplo 4.12

Encuentre el determinante de la matriz:

$$A = \begin{bmatrix} 2 & 2 & 0 & 4 \\ 3 & 3 & 2 & 2 \\ 0 & 1 & 3 & 2 \\ 2 & 0 & 2 & 1 \end{bmatrix}$$

Solución

$$\begin{vmatrix} 2 & 2 & 0 & 4 \\ 3 & 3 & 2 & 2 \\ 0 & 1 & 3 & 2 \\ 2 & 0 & 2 & 1 \end{vmatrix} = 2\begin{vmatrix} 1 & 1 & 0 & 2 \\ 3 & 3 & 2 & 2 \\ 0 & 1 & 3 & 2 \\ 2 & 0 & 2 & 1 \end{vmatrix}$$

Propiedad de multiplicación por un escalar

$$= 2\begin{vmatrix} 1 & 1 & 0 & 2 \\ 3 & 3 & 2 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & -2 & 2 & -3 \end{vmatrix}$$

El 1er renglón es multiplicado por –2 y se suma al 4º renglón.

$$= 2\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 2 & -4 \\ 0 & 1 & 3 & 2 \\ 0 & -2 & 2 & -3 \end{vmatrix}$$

Se multiplica el 1er renglón por –3 y se suma al 2º.

$$= 2\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 2 & -4 \\ 0 & -2 & 2 & -3 \end{vmatrix}$$

Se intercambian renglones.

$$= 2\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 2 & -4 \\ 0 & 0 & 8 & 1 \end{vmatrix}$$

Se multiplica el 2º renglón y se suma al 4º.

$$=(-2)(2)\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 8 & 1 \end{vmatrix}$$

Propiedad de multiplicación por un escalar.

$$=(-2)(2)\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 17 \end{vmatrix}$$

Se multiplica el 3º renglón por –8 y se suma al 4º.

De aquí $\det A = (-2)(2)(17) = -68$

El ejemplo anterior ilustra que si una matriz A de orden $n \times n$ se puede reducir a una matriz triangular superior U con n pivotes (distintos de cero), entonces det (A) es igual $a \pm$ el producto de los pivotes. Si en algún punto de la reducción no se puede hallar un pivote distinto de cero, entonces el determinante de esta matriz y el determinante de A son igual a cero.

Por ejemplo:
Sea A matriz de 5×5 que se ha reducido y se tiene:

$$\begin{bmatrix} p_1 & \times & \times & \times & \times \\ 0 & p_2 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

Desarrollando repetidamente por menores en las primeras columnas vemos fácilmente que el determinante es cero.

Teorema 4.3: Criterio de invertibilidad para una matriz cuadrada
Sea A una matriz de $n \times n$, entonces A tiene inversa si y solo si det $A \neq 0$.

Demostración
→) Sí A es invertible entonces $I = AA^{-1}$, por lo que $1 = \det(I) = \det(A)\det(A^{-1})$; por lo tanto, det $A \neq 0$.

←) Se tiene det $A \neq 0$, entonces se pueden aplicar operaciones elementales en A para expresarla como una matriz equivalente U, esto es:

$$E_k \cdots\cdots E_2\, E_1\, A = U \quad o \quad A = E_1^{-1}\, E_2^{-1} \cdots\cdots E_k^{-1}\, U$$

De aquí que:
$$\det A = \det(E_1^{-1}) \det(E_2^{-1}) \ldots\ldots \det(E_k^{-1}) \det U$$

Como det $A \neq 0$, entonces det $(U) \neq 0$, de donde U no tiene renglones iguales a cero y por lo tanto $U = I$.

TEOREMA 4.4
Para una matriz A de $n \times n$ las siguientes afirmaciones son equivalentes:

a) La ecuación $Ax=0$ tiene la única solución $x=0$.
b) El sistema lineal $Ax=b$ tiene una solución única para cualquier vector columna b n-dimensional.
c) La matriz A es no-singular, es decir A^{-1} existe.
d) El determinante de A es distinto de cero.
e) El algoritmo de eliminación gaussiana se puede aplicar al sistema lineal $Ax=b$.

Ya hemos visto el método de desarrollo por cofactores para encontrar un determinante, ahora encontraremos una fórmula para calcular la inversa de una matriz.

Definición 4.3
Si A es una matriz de $n \times n$ y a'_{ij} es el cofactor de a_{ij} entonces la matriz:

$$A' = \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a'_{n1} & a'_{n2} & \cdots & a'_{nn} \end{bmatrix}$$

Se llama la **matriz de cofactores de *A***. La transpuesta de esta matriz se denomina la **adjunta de *A*** y se denota por adj(A).

Ejemplo 4.13
Sea la matriz:

Los cofactores de A son:

$$\begin{array}{lll} a'_{11} = 12 & a'_{12} = 6 & a'_{13} = -16 \\ a'_{21} = -4 & a'_{22} = 2 & a'_{23} = 16 \\ a'_{31} = 12 & a'_{32} = -10 & a'_{33} = 16 \end{array}$$

Por lo que la matriz de cofactores es:

$$A'=\begin{bmatrix} 12 & 6 & -16 \\ -4 & 2 & 16 \\ 12 & -10 & 16 \end{bmatrix} \text{ y la adjunta de } A \text{ es: } \quad adj\ \mathrm{A}=\begin{bmatrix} 12 & -4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix}$$

Ahora es posible establecer una fórmula para calcular la inversa de una matriz invertible.

TEOREMA 4.5

Si A es una matriz invertible, entonces:

$$A^{-1} = \frac{1}{\det A}\ adj\ (A)$$

Demostración

Primero se demostrará que:

$$A \text{ adj}(A) = \det(A)I$$

Considere el producto:

$$A \text{ adj}(A) = \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{11} & a'_{12} & \cdots & a'_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{11} & a'_{12} & \cdots & a'_{1n} \end{bmatrix} \begin{bmatrix} a'_{11} & a'_{21} & a'_{j1} & \cdots & a'_{n1} \\ a'_{11} & a'_{21} & a'_{j1} & \cdots & a'_{n1} \\ \vdots & \vdots & \vdots & & \vdots \\ a'_{11} & a'_{21} & a'_{j1} & \cdots & a'_{n1} \end{bmatrix}$$

El elemento que se encuentra en el *i-ésimo* renglón y en la j-ésima columna de A adj (A) es:

$$a_{i1}\, a'_{j1} + a_{i2}\, a'_{j2} + \cdots + a_{in}\, a'_{jn} \qquad (4.6)$$

Si $i=j$ entonces (4.6) es el desarrollo por cofactores de $\det(A)$ a lo largo del i-ésimo renglón de A.

Si $i \neq j$ entonces los elementos y los cofactores provienen de diferentes renglones de A, por lo que el valor de (4.6) es cero.

Note que:

$$a_{21}\, a_{11} + a_{22}\, a'_{12} + \cdots\cdots + a_{2n}\, a'_{2n} = \begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = 0$$

Por lo tanto:

$$\text{A adj}(A) = \begin{bmatrix} \det \text{A} & & \\ & \det \text{A} & \\ & & \det \text{A} \end{bmatrix} = (\det A)^{\text{I}} \qquad (4.7)$$

Como A es invertible det $(A) \neq 0$, entonces la ecuación (4.7) se puede reescribir como:

$$\frac{1}{\det A} = [A \text{ adj } A] = I$$

Entonces:

$$A = \left[\frac{1}{\det A}\right] = \text{adj}(A) = I$$

$$A^{-1} = \frac{1}{\det A} = \text{adj}(A)$$

Ejemplo 4.14

Encuentre la inversa de la matriz: $A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 2 & 0 \\ 3 & 1 & 1 \end{bmatrix}$

Usando la fórmula para la inversa.

Solución

Primero se verifica que el determinante de A sea diferente de cero: como el det(A) =4, entonces A es invertible. Para calcular la adjunta de A tenemos que los cofactores de A son:

$a'_{11} = 2$ $\quad a'_{12} = -2$ $\quad a'_{13} = -4$

$a'_{21} = 1$ $\quad a'_{22} = 1$ $\quad a'_{23} = -4$

$a'_{31} = -2$ $\quad a'_{32} = 2$ $\quad a'_{33} = 8$

Por lo que la matriz de cofactores es:

$$A' = \begin{bmatrix} 2 & -2 & -4 \\ 1 & 1 & -4 \\ -2 & 2 & 8 \end{bmatrix} \quad \text{Y la adjunta de } A \text{ es:} \qquad adj\ A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 1 & -4 \\ -4 & -4 & 8 \end{bmatrix}$$

De aquí que:

$$A' = \frac{\text{adj } A}{\det A} = \begin{bmatrix} ½ & ¼ & -½ \\ -½ & ¼ & ¼ \\ -1 & -1 & 2 \end{bmatrix}$$

Para matrices de orden mayor que 3×3 este método para invertir matrices es operacionalmente inferior a la técnica de Gauss-Jordan.

De la misma manera, con frecuencia es útil tener una fórmula en términos de determinantes para la solución de un sistema de ecuaciones $Ax=b$, donde A es una matriz invertible, por lo cual enunciamos el siguiente teorema.

TEOREMA 4.6: REGLA DE CRAMER

Si $Ax=b$ es un sistema de n ecuaciones lineales en n incógnitas tal que $\det(A)\neq 0$, entonces el sistema tiene solución única y está dada por:

$$x_k = \frac{\det B_k}{\det A} \qquad \text{para } k = 1, 2, \cdots, n \tag{4.8}$$

Donde B_k es la matriz que se obtiene al reemplazar los elementos de la j-ésima columna de A por los elementos del vector columna b.

Demostración

Como A es invertible, la solución única del sistema $Ax=b$ se puede expresar como:

$$x_k = A^{-1}\, b = \frac{\det A}{\det A}\ b$$

Comparando con (1), vemos que solo se necesita demostrar que la k-ésima componente del vector columna adj $A\ b$ está dado por el determinante de la matriz:

$$B_k = \begin{bmatrix} a_{11} & b_1 & \cdots & a_{ln} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{bmatrix} = 0$$

obtenida de A mediante el reemplazo de la k-ésima columna por b; esto es, el reemplazo de a_{jk} por b_j. Al desarrollar det B_k por menores en la k-ésima columna se tiene:

$$\det B_k = \sum_{j=1}^{n} a'_{jk}\, b_j$$

Pero ésta es la k-ésima componente de:

$$(\text{adj } A)\, b = \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{n1} \\ a'_{1k} & a'_{2k} & \cdots & a'_{nk} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{1n} & a'_{2n} & \cdots & a'_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= a'_{1k}\, b_1 + \cdots + a'_{nk}\, b_n$$

Lo que completa la demostración.

Ejemplo 4.15

Resuelva el siguiente sistema lineal usando la regla de Cramer.

$$\begin{aligned} x_1 + \quad\;\; + 2x_3 &= 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solución

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix} \quad \det A = 44$$

$$B_1 = \begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{bmatrix} \qquad B_2 = \begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{bmatrix} \qquad B_3 = \begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix}$$

$$\det B_1 = -40 \qquad \det B_2 = 72 \qquad \det B_3 = 152$$

Por lo tanto:

$$x_1 = \frac{\det(B_1)}{\det(A)} = \frac{-40}{44} = \frac{-10}{11}$$

$$x_2 = \frac{\det(B_2)}{\det(A)} = \frac{72}{44} = \frac{-18}{11}$$

$$x_3 = \frac{\det(B_3)}{\det(A)} = \frac{152}{44} = \frac{38}{11}$$

4.7. RESUMEN

1. Un determinante de segundo orden está definido por:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

2. El área del paralelogramo con vértice en el origen determinado por los vectores distintos de cero *a* y *b* de R^2, es el valor absoluto del determinante de la matriz que tiene vectores fila *a* y *b*.

3. El producto cruz de los vectores *b* y *c* de R^3 se puede calcular usando el determinante simbólico:

$$\begin{bmatrix} i & j & k \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_{31} \end{bmatrix}$$

   Este producto cruz *bxc* es perpendicular a: *b* y *c*.

4. El área del paralelogramo determinado por los vectores distintos de cero *b* y *c* de R^3es $\|b \times c\|$.

5. El volumen de la caja determinada por los vectores distintos de cero *a, b* y *c* de R^3 es el valor absoluto del determinante de la matriz que tiene vectores fila *a, b* y *c.* Este determinante también es igual a: $a(b\times c)$.

6. El cofactor de un elemento a_{ij} en una matriz cuadrada *A* es $(-1)^{i+j}\,|A_{ij}|$, donde A_{ij} es la matriz obtenida de *A* al eliminar la *i*-ésima fila y la *j*-ésima columna.

7. El determinante de una matriz de $n\times n$ se puede definir inductivamente desarrollando por menores en la primera fila. El determinante se puede calcular desarrollando por menores usando cualquier fila o columna; es la suma de los productos de los registros en esa fila o columna con los cofactores de los registros. En general, dicho cálculo es muy largo.

8. Las operaciones elementales en filas tienen el efecto siguiente en el determinante de una matriz cuadrada *A*:

   a) Si se intercambian dos filas diferentes de *A*; cambia el signo del determinante.
   b) Si se multiplica una sola fila de *A* por un escalar, el determinante se multiplica por el escalar.
   c) Si un múltiplo de una fila se suma a una fila diferente, el determinante no cambia.

9. La inversa de una matriz invertible *A* está dada por la fórmula:

$$A^{-1} = \frac{1}{\det(A)}\, adj\,(A)$$

10. Si A es invertible, entonces un sistema de ecuaciones $Ax=b$ tiene la solución única x, cuya *k*-ésima componente está dada explícitamente por la fórmula

$$x_k = \frac{\det B_k}{\det(A)}$$

Donde la matriz B_k se obtiene de *A* reemplazando la *k*-ésima columna de *A* por *b*.

4.8. APLICACIONES DE LOS DETERMINANTES

(Tomado de *Aplicaciones de álgebra lineal,* Rorres y Anton 1977)

Los determinantes tienen muchas aplicaciones en la ingeniería civil sobre todo en el área de estructuras, en ingeniería eléctrica y en ingeniería industrial por mencionar algunas, aquí se presentan algunas aplicaciones publicadas en el libro de Rorres y Anton.

ECUACIONES DE CURVAS Y SUPERFICIES QUE PASAN POR PUNTOS ESPECÍFICOS

El siguiente enunciado es uno de los postulados fundamentales de la teoría del álgebra lineal:

Se mostrará cómo este postulado sirve para determinar las ecuaciones de las curvas y superficies que pasan por puntos dados, desarrollando algunos ejemplos específicos.

> *La solución de un sistema lineal homogéneo con un número de ecuaciones igual al número de incógnitas no es trivial si, y solo si, el determinante del sistema es cero.*

LA RECTA QUE PASA POR DOS PUNTOS

Si se tienen como datos (los puntos diferentes del plano, (x_1y_1) y (x_2y_2), hay solo una recta:

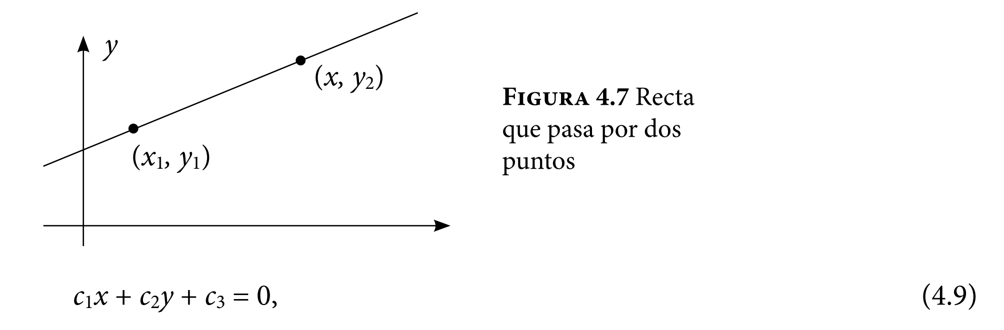


Figura 4.7 Recta que pasa por dos puntos

$$c_1x + c_2y + c_3 = 0, \tag{4.9}$$

que pasa por ellos. Observe que no pueden ser cero simultáneamente y que son únicos para una recta dada. Como pertenecen a la recta, al sustituirlos en (4.9) se obtienen las ecuaciones

$$c_1x_1 + c_2y_1 + c_3 = 0 \tag{4.10}$$
$$c_1x_2 + c_2y_2 + c_3 = 0 \tag{4.11}$$

Las ecuaciones (4.9), (4.10) y (4.11), pueden agruparse formando un sistema:

$$xc_1 + yc_2 + c_3 = 0$$
$$x_1c_1 + y_1c_2 + c_3 = 0$$
$$xc_1 + y_2c_2 + c_3 = 0$$

Que es un sistema homogéneo de tres ecuaciones para c_1, c_2 y c_3 y como estos coeficientes no pueden ser todos cero, el sistema tiene una solución que no es trivial. Así, el determinante del sistema debe ser, necesariamente igual a cero. Entonces:

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0 \tag{4.12}$$

En consecuencia, todo punto (x,y) de la recta satisface (4.12), recíprocamente, todo punto (x,y) que satisface a (4.12) pertenece a la recta.

Ejemplo 4.16

Obtener la ecuación de la recta que pasa por los puntos (2,1) y (3,7).

Solución

Al sustituir las coordenadas de los puntos dados en la ecuación (4.12) se obtiene:

$$\begin{vmatrix} x & y & 1 \\ 2 & 1 & 1 \\ 3 & 7 & 1 \end{vmatrix} = 0$$

Desarrollando por cofactores tomando como base el primer renglón, se llega a la ecuación:

$$-6x + y + 11 = 0$$

CIRCUNFERENCIA QUE PASA POR TRES PUNTOS

Si se tienen como datos tres puntos del plano, diferentes y no alineados, (x_1, y_1), (x_2, y_2) y (x_3, y_3) de la geometría analítica se sabe que solo existe una circunferencia, dada por la ecuación:

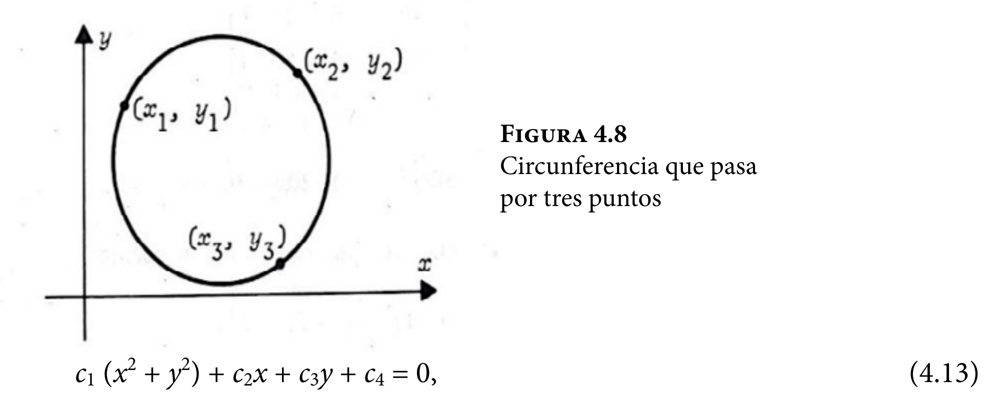


Figura 4.8
Circunferencia que pasa por tres puntos

$$c_1\,(x^2 + y^2) + c_2 x + c_3 y + c_4 = 0, \tag{4.13}$$

Que pasa por estos tres puntos. Sustituyendo sus coordenadas en la ecuación, se obtienen:

$$c_1\,(x_2^2 + y_2^2) + c_2 x_1 + c_3 y_1 + c_4 = 0 \tag{4.14}$$
$$c_1\,(x_2^2 + y_2^2) + c_2 x_2 + c_3 y_2 + c_4 = 0 \tag{4.15}$$
$$c_1\,(x_3^2 + y_3^2) + c_2 x_3 + c_3 y_3 + c_4 = 0 \tag{4.16}$$

Como en el caso anterior, las ecuaciones (4.13) a (4.16) forman un sistema lineal homogéneo para cuya solución no es la trivial. En consecuencia, el determinante del sistema es igual a cero:

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ x_1^2+y_1^2 & x_1 & y_1 & 1 \\ x_2^2+y_2^2 & x_2 & y_2 & 1 \\ x_3^2+y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0 \tag{4.17}$$

Esta es la forma del determinante para la ecuación de una circunferencia.

Ejemplo 4.17
Obtener la ecuación de la circunferencia que pasa por los puntos (1, 7), (6, 2) y (4, 6).

Solución

Sustituyendo las coordenadas de los puntos en la ecuación (4.17), se obtiene el determinante

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ 50 & 1 & 7 & 1 \\ 40 & 6 & 2 & 1 \\ 52 & 4 & 6 & 1 \end{vmatrix} = 0,$$

que se reduce a:

$$10\,(x^2+y^2)\ -20x\ 40y\ -200 = 0$$

Esta expresión se puede escribir en la forma ordinaria:

$$(x-1)^2 + (y-2)^2 = 5^2$$

Así, se trata de una circunferencia cuyo centro es el punto (1, 2) y su radio es 5.

SECCIÓN CÓNICA QUE PASA POR CINCO PUNTOS

La forma general de la ecuación de una sección cónica contenida en el plano (una parábola, una hipérbola o una elipse y las formas degeneradas correspondientes) es:

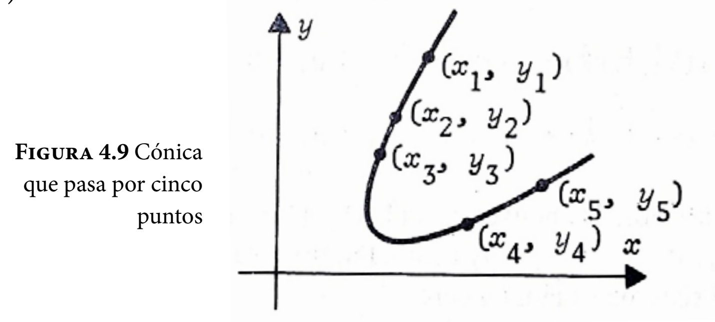


Figura 4.9 Cónica que pasa por cinco puntos

$$c_1x^2 + c_2\,xy + c_3y^2 + c_4x + c_5y + c_6 = 0$$

Esta forma tiene seis coeficientes, de los cuales solo se necesitan cinco, ya que la expresión puede dividirse entre cualquiera de ellos que no sea cero. Así, solo se tienen que determinar cinco coeficientes, por lo que bastan cinco puntos distintos del plano para determinar la ecuación de la sección cónica (Figura 4.9).

Como en los ejemplos anteriores, la ecuación puede escribirse en la forma del determinante:

$$\begin{vmatrix} x^2 & xy & y^2 & x & y & 1 \\ x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \end{vmatrix} = 0 \qquad (4.18)$$

Ejemplo 4.18

Un astrónomo quiere determinar la órbita de un asteroide alrededor del Sol. Para ello, introduce un sistema de coordenadas cartesianas en el plano de la órbita con el Sol en el origen. Las unidades de medición utilizadas en los ejes son unidades astronómicas (1 unidad astronómica = la distancia media de la Tierra al Sol = 93 millones de millas). Por la primera ley de Kepler, se sabe que la órbita es una elipse. En consecuencia, con cinco observaciones del asteroide en horas diferentes, determina cinco puntos de la órbita cuyas coordenadas son:

(5.764,0.648), (6.286,1.202), (6.759,1.823), (7.168,2.526), (7.480,3.360)

Obtener la ecuación de la órbita.

Solución

Sustituyendo las coordenadas de los cinco puntos en la ecuación (4.18), se obtiene:

$$\begin{vmatrix} x^2 & xy & y^2 & x & y & 1 \\ 33.224 & 3.735 & 0.420 & 5.764 & 0.648 & 1 \\ 39.514 & 7.556 & 1.445 & 6.286 & 1.202 & 1 \\ 45.684 & 12.322 & 3.323 & 6.759 & 1.823 & 1 \\ 51.380 & 18.106 & 6.381 & 7.168 & 2.526 & 1 \\ 55.950 & 25.133 & 11.290 & 7.480 & 3.360 & 1 \end{vmatrix} = 0$$

Desarrollando por cofactores en base al primer renglón, el determinante se reduce a:

$$x^2 - 1.04xy + 1.30\,y^2 - 3.90x - 2.93y\ 5.49 = 0$$

La figura 4.10 es un diagrama de la órbita donde aparecen los cinco puntos dados.

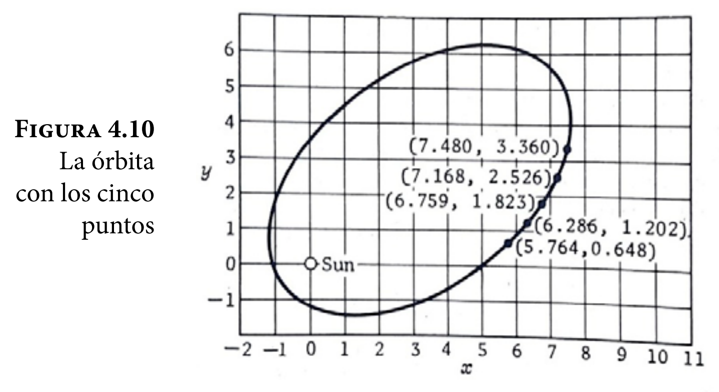


Figura 4.10 La órbita con los cinco puntos

PLANO QUE PASA POR TRES PUNTOS

En el ejercicio 6 se le pide al lector demostrar que el plano cuya ecuación en el espacio de tres dimensiones es:

$$c_1x + c_2y + c_3z + c_4 = 0$$

Y que pasa por tres puntos no colineales (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), puede también expresarse por una ecuación en forma de determinante:

$$\begin{bmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{bmatrix} = 0 \qquad (4.19)$$

Ejemplo 4.19

La ecuación del plano que pasa por los puntos no colineales (1, 1, 0), (2, 0, –1) y (2, 9, 2) es:

$$\begin{bmatrix} x & y & z & 1 \\ 1 & 1 & 0 & 1 \\ 2 & 0 & -1 & 1 \\ 2 & 9 & 2 & 1 \end{bmatrix} = 0$$

Que se reduce a:

$$2x - y + 3z - 1 = 0.$$

ESFERA QUE PASA POR CUATRO PUNTOS

En el ejercicio 7 se le pide al lector que demuestre que la esfera cuya ecuación en el espacio de tres dimensiones es:

$$c_1 (x_2 + y^2 + z^2) + c_2 x + c_3 y + c_4 z + c_5 = 0$$

Y que pasa por cuatro puntos no coplanares (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4), está también dada por la ecuación en forma de determinante:

$$\begin{vmatrix} x^2 + y^2 + z^2 & x & y & z & 1 \\ x_1^2 + y_1^2 + z_1^2 & x_1 & y_1 & z_1 & 1 \\ x_2^2 + y_2^2 + z_2^2 & x_2 & y_2 & z_2 & 1 \\ x_3^2 + y_3^2 + z_3^2 & x_3 & y_3 & z_3 & 1 \\ x_4^2 + y_4^2 + z_4^2 & x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0 \qquad (4.20)$$

Ejemplo 4.20

La ecuación de la esfera que pasa por los puntos (0, 3, 2), (1,–1, 1), (2, 1, 0) y (–1, 1, 3) es:

$$\begin{vmatrix} x^2 + y^2 + z^2 & x & y & z & 1 \\ 13 & 0 & 3 & 2 & 1 \\ 3 & 1 & -1 & 1 & 1 \\ 5 & 2 & 1 & 0 & 1 \\ 11 & -1 & 1 & 3 & 1 \end{vmatrix} = 0$$

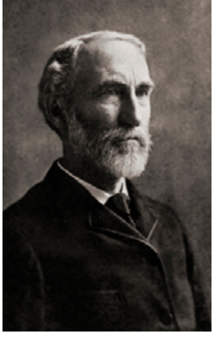
Que se reduce a:

$$x^2 + y^2 + z^2 - 4x - 2y - 6z + 5 = 0,$$

Y que en la forma ordinaria es:

$$(x-2)^2 + (y-1)^2 + (z-3)^2 = 9$$

4.9. NOTAS HISTÓRICAS



La noción de producto cruz surgió de los intentos de Sir William Rowan Hamilton para desarrollar una multiplicación para "ternas"; esto es, vectores de R^3. La simbología actual para el producto cruz apareció por primea vez en el breve texto "Elements of Vector Análisis" (1881) del norteamericano Josiah Willard Gibbs (1839-1903), profesor de física matemática en Yale, que lo escribió para sus cursos de electricidad y magnetismo de Yale, y de mecánica en Johns Hopkins.

La interpretación como volumen de un determinante apareció por primera vez en 1773, en un artículo sobre mecánica de Joseph Louis Lagrange (1736-1813). Lagrange observó que si los puntos M, M', M'' tienen coordenadas (x, y, z), $(x'\, y''\, z')$ y (x'', y'', z'') respectivamente, entonces el tetraedro con vértices en el origen y esos tres puntos tendrá volumen:



$$(1/6)\left[z(x'y''-y'xx'') + z'(yx''-xy'')+ z''(xy'-yx')\right]$$

Esto es:

$$(1/6)\ \det \begin{bmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{bmatrix}$$

Lagrange nació en Turín y pasó la mayor parte de su carrera matemática en Berlín y Paris; aportó resultados importantes en temas tan variados como cálculo de variaciones, mecánica celeste, teoría de números y teoría de ecuaciones. Entre sus trabajos más famosos están el Tratado de Mecánica Analítica (1788), en el cual presentó los distintos principios de la mecánica desde un solo punto de vista, y la Teoría de Funciones Analíticas (1797), donde intentó basar el cálculo diferencial en la teoría de series de potencias.

La teoría de los determinantes se desarrolló a partir de los esfuerzos de muchos matemáticos de finales del siglo XVIII y principios del XIX. Además de Gabriel Cramer (1704-1752), Etienne Bezout (1739-1783), en 1764, y Alexandre- Theophile Vandermonde (1735-1796), en 1771, dieron varios métodos para



calcular determinantes. En un trabajo sobre cálculo integral, Pierre Simon Laplace (1749-1827) trató sistemas de ecuaciones lineales repitiendo el trabajo de Cramer, pero también enunció y probó la regla de que el intercambio de dos columnas adyacentes del determinante cambia el signo y mostró que un determinante con dos columnas iguales es cero.

El más completo de los primeros trabajos sobre determinantes es el de Augustin-Louis Cauchy (1789-1857) en 1812. En éste trabajo, Cauchy introdujo el nombre de **"determinante"** para reemplazar varios términos antiguos, usó la notación hoy en boga del doble subíndice para un arreglo cuadrado de números, definió el arreglo de adjuntos (o menores) a un arreglo dado, y mostró que podemos calcular el determinante desarrollando en cualquier fila o columna. Además, Cauchy probó de nuevo muchos teoremas estándar sobre determinantes más o menos conocidos durante los 50 años anteriores.



Cauchy fue el matemático más prolífico del siglo XIX, hizo aportaciones en áreas como análisis complejo, cálculo, ecuaciones diferenciales y mecánica. En particular, escribió el primer texto de cálculo usando el enfoque moderno de ε-δ para la continuidad. Políticamente fue conservador; cuando la Revolución de julio de 1830 reemplazó al rey Borbón Carlos X por el rey Orleáns Luis Felipe Cauchy se negó a prestar el juramento de fidelidad, perdiendo así sus puestos en la Ecole Polytechnique y en el College de France y se exilió en Turín y Praga.

La primera aparición del determinante de una matriz cuadrada en Europa occidental ocurrió en 1683, en una carta de Gottfried von Leibniz al marqués de L'Hopital. Leibniz escribe un sistema de tres ecuaciones con dos incógnitas con coeficientes "numéricos" abstractos como sigue:



$$10 + 11x + 12y = 0$$
$$20 + 21x + 22y = 0$$
$$30 + 31x + 32y = 0$$

En donde se observa que cada número de coeficiente tenía "dos caracteres: el primero que marca en qué ecuación se presenta y el segundo marca a

que letra pertenece". Después, procedió a eliminar, y luego x.
Para mostrar que el criterio para que el sistema de ecuaciones tenga solución es:

$$(10)(21)(32) + (11)(22)(30) + (12)(20)(31) = (10)(22)(31) + (11)(20)(32) + (12)(21)(30).$$

Es evidente que esto equivale a nuestra condición de que el determinante de la matriz de coeficientes se debe anular. Lamentablemente, la carta no se publicó hasta 1850, por lo que no influyó en trabajos posteriores.

Los determinantes también aparecieron en el trabajo contemporáneo del matemático japonés Takakazu Seki (1642-1708).



La regla de Cramer apareció por primera vez en toda su generalidad en una obra de Gabriel Cramer (1704-1752) titulada "Introducción al análisis de curvas algebraicas" (1750). El problema en el que Cramer estaba interesado era el de determinar la ecuación de una curva plana de grado dado que pasara por cierto número de puntos dados. Formuló el teorema de que una curva de grado n-ésimo está determinada cuando se conocen (½)*n* (*n*+3) puntos de la curva. Por ejemplo, una curva de segundo grado que él escribía como:

$$A + By + Cx + Dy^2 + Exy + x^2 = 0$$

Está determinada por cinco puntos. Entonces, la cuestión es cómo determinar A, B, C, D, E dados los cinco puntos. El método obvio es sustituir sucesivamente en la ecuación (*) las coordenadas de cada uno de los cinco puntos. Esto da cinco ecuaciones para los coeficientes desconocidos. Cramer cita el apéndice del folleto, en donde da su regla general: "se halla el valor de cada incógnita formando n fracciones cuyo común denominador tiene tantos términos como permutaciones de *n* objetos hay" y pasa a explicar exactamente cómo se calculan estos términos como productos de ciertos coeficientes de las n ecuaciones, cómo se determina el signo apropiado para cada término y cómo se determinan los n numeradores de las fracciones al reemplazar ciertos coeficientes de este cálculo por los términos constantes del sistema.

Cramer fue un matemático suizo que enseñó en la Académie de Calvin, en Ginebra, desde 1724 hasta su muerte. Fue un gran sabio, pero su trabajo se vio oscurecido por contemporáneos tan brillantes como Euler y los Bernoulli.

1. http://tochtli.fisica.uson.mx/electro/vectores/definici%C3%B3n_de_vectores.htm

4.10. EJERCICIOS PROPUESTOS

1. Considere la matriz:

$$A=\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}$$

Cuyos valores son tales que el det A=15. Calcule el det B y det C si:

$$B=\begin{bmatrix} a & e & i & m \\ b & f & j & n \\ c & g & k & o \\ d & h & l & p \end{bmatrix} \qquad C=\begin{bmatrix} 2a & 2e & 2m & 2i \\ b & f & n & j \\ 3c & 3g & 3o & 3k \\ 5d & 6h & 6p & 5l \end{bmatrix}$$

2. Obtenga el determinante de A si:

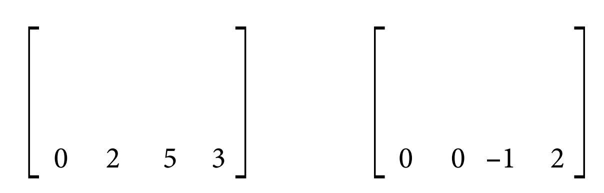
$$A=\begin{bmatrix} 3 & 2 & 2 & 2 & 2 \\ 2 & 3 & 2 & 2 & 2 \\ 2 & 2 & 3 & 2 & 2 \\ 2 & 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 2 & 3 \end{bmatrix}$$

3. Calcule el determinante de la matriz:

$$A=\begin{bmatrix} x+a_1 & a_2 & a_3 & a_4 & a_5 \\ a_1 & x+a_2 & a_3 & a_4 & a_5 \\ a_1 & a_2 & x+a_3 & a_4 & a_5 \\ a_1 & a_2 & a_3 & x+a_4 & a_5 \\ a_1 & a_2 & a_3 & a_4 & x+a_5 \end{bmatrix}$$

4. Mediante operaciones en filas encuentre una triangular superior U y calcule el determinante de las siguientes matrices:

$$A=\begin{matrix} 1 & 2 & -2 & 0 \\ 2 & 3 & -4 & 1 \\ -1 & -2 & 0 & 2 \end{matrix} \qquad B=\begin{matrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \end{matrix}$$



5.

a) Encuentre la solución del sistema

$2x_1+3x_2-x_3= 4$
$x_1-2x_2+x_3= 6$
$x1-12x_2+5x_3= 10$
Usando la regla de Cramer.

b) Demuestre que el sistema lineal:

$2x_1+3x_2-x_3= 4$
$x_1-2x_2+x_3= 6$
$-x1-12x_2+5x_3= 9$
No tiene solución. Calcule B_1, B_2 y B_3.

c) Demuestre que el sistema lineal:

$2x_1+3x_2-x_3= 4$
$x_1-2x_2+x_3= 6$
$-x1-12x_2+5x_3= 1$
Tiene un número infinito de soluciones. Calcule B_1, B_2 y B_3.

6. Demostrar que la ecuación (4.19) es la ecuación, en el espacio tridimensional, de un plano que pasa por tres puntos no colineales.

7. Demostrar que la ecuación (4.20) es la ecuación, en el espacio tridimensional, de una esfera que pasa por cuatro puntos no coplanares.

Capítulo 5
ESPACIOS VECTORIALES

5.1. INTRODUCCIÓN

Hemos visto cómo se puede efectuar la eliminación para simplificar un sistema lineal $Ax=b$; afortunadamente, no solo se facilita el cálculo del valor de x, sino que también se responde a las cuestiones teóricas acerca de su existencia y su unicidad. El objetivo fundamental de esta parte es adquirir un conocimiento diferente y más profundo del problema.

Para ello necesitamos el concepto de espacio vectorial, introducimos la idea con los espacios más importantes R, R^2, $R^3, \cdots$ El espacio R^2 se representa por el ya conocido plano x–y, donde las dos componentes del vector son las coordenadas x, y del punto correspondiente. También es familiar R^3 con las tres componentes que dan un punto en el espacio tridimensional y R es una recta. Lo importante en álgebra lineal es que la extensión a n dimensiones es directa.

En estos espacios, como en todos los espacios vectoriales, son posibles dos operaciones: la suma de dos vectores cualesquiera y la multiplicación de escalares por vectores.

5.2. ESPACIOS VECTORIALES Y ALGEBRA EN R^N

Si bien se van a dar las definiciones para espacios de menor dimensión, en esta sección se define el espacio en R^n dado que es el caso general y posteriormente se presenta la definición de un espacio vectorial en R.

Sean u y v vectores en R^n, donde $u = (u_1, u_2, \cdots, u_n)$ y $v = (v_1, v_2, \cdots, v_n)$

La suma $u + v$ es el vector: $u+v = (u_1+v_1, u_2+v_2, \cdots, u_n+v_n)$

Si r es un escalar, el múltiplo escalar ru se define:

$$ru = (ru_1, ru_2, \cdots, ru_n)$$

Si u es un vector en R^n se define el negativo (o inverso aditivo) de u mediante $-u$ y se define como:

$$-u = (-u_1, -u_2, \cdots, -u_n)$$

La sustracción de vectores en R^n se define como $v-u = v+(-u)$ o en términos de las componentes:

$$v-u = (v_1, v_2, \cdots, v_n) + (-u_1, -u_2, \cdots, -u_n) = v_1-u_1, v_2-u_2, \cdots, v_n-u_n$$

Nota: Para sumar dos vectores, ambos deben estar contenidos en el mismo espacio R^n. La suma de vectores con diferente número de componentes no está definida.

Las propiedades más importantes de la suma y el producto por un escalar en R^n se enumeran en el siguiente teorema.

Teorema 5.1

Sean u, v y w elementos de R^n y r, s escalares, entonces:

a) $u+v = v+u$
b) $u+(v+w) = (u+v)+w$
c) $u+0 = 0+u = u$
d) $u+(-u) = 0$ *es decir* $u-u = 0$
e) $r(su) = (rs)u$
f) $r(u+v) = ru+rv$
g) $(r+s)u = ru+su$
h) $1u = u$

Este teorema permite manejar los vectores en R^n sin tener que expresarlos en términos de sus componentes, exactamente como se trabaja con los números reales.

Para extender las nociones de distancia, norma y ángulo a R^n se comenzará con la siguiente generalización de producto punto en R^2 y en R^3.

Definición 5.1

Si u y v son vectores en R^n, entonces el producto interior euclidiano se define como:

$$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Los vectores u y v se dicen ortogonales (o perpendiculares) si su producto interior es igual a cero, es decir, si $u \cdot v = 0$.

Ejemplo 5.1

Encuentre el producto interior euclidiano de los vectores u y v en R^4 donde u y v son:

$u = (-1, 3, 5, 7) \qquad v = (5, -4, 7, 0)$

Solución

$\mathrm{u} \cdot \mathrm{v} = (-1)(5) + (3)(-4) + (5)(7) + (7)(0) = 18$

Teorema 5.2

Si u, v y w son vectores en R^n y r es un escalar, entonces:

a) $u \cdot v = v \cdot u$
b) $(u + v) \cdot w = u \cdot w + v \cdot w$
c) $r(u \cdot v) = (ru) \cdot v = u \cdot (rv)$
d) $u \cdot u\ ^{3} 0$ *y* $u \cdot = 0$ si y solo si $u = 0$

Demostración del inciso d:

$u \cdot u = u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$ además la igualdad se cumple si y solo si:
$u_1 = u_2 = \cdots = u_n$, es decir solo si $u = 0$

***Definición 5.2* Norma y distancia en *R^n*.**

Dados los vectores u y v en R^n se define la distancia $d(u,v)$ entre los puntos u y v como:

$$d\,(u,v) = \sqrt{(u_1 - v_1)^2 + (u_2 + v_2)^2 + \cdots\cdots + (u_n + v_n)^2}$$

La norma o longitud del vector u denotada $\|u\|$ es definida como:

$$\|u\| = u \cdot u = \sqrt{u_1^{\,2} + u_2^{\,2} + \cdots\cdots + u_n^{\,2}}$$

Donde $u \cdot u \geq 0$ y entonces la raíz cuadrada existe. Observe que:

$$d\,(u,v) = \|u - v\|$$

Es importante tener una buena apreciación del concepto de norma, dada su relación con otras ideas del álgebra lineal. Una primera relación de interés es la que se observa entre norma y distancia. Considere los puntos $P_1 = (a,\, b)$ y $P_2 = (c,\, d)$ en R^2.

$$\|P_1\| = \sqrt{a^2 + b^2} \qquad d\,(P_1, P_2) = \sqrt{(a-c)^2 + (b-d)^2}$$

Así $\|P_1\|$ corresponde a la distancia del origen al punto P_1 y $d\,(P_1, P_2)$ corresponde a la distancia entre los puntos P_1 y P_2.

Ejemplo 5.2

Dados los vectores $u = (1, 2, 0, 3)$ y $v = (0, 3, 5, 2)$ encuentre la norma de u, la norma de v y la distancia entre ellos.

Solución

$$\|u\| = \sqrt{(1)^2 + (2)^2 + (0)^2 + (3)^2} = \sqrt{14}$$

$$\|v\| = \sqrt{(0)^2 + (3)^2 + (5)^2 + (2)^2} = \sqrt{38}$$

$$D(u,v) = \sqrt{(1-0)^2 + (2-3)^2 + (0-5)^2 + (3-2)^2} = \sqrt{28}$$

Propiedades de la Magnitud o Norma

a) $|v|^2 = v$

b) $|rv| = |r|\,|v|$

Teorema 5.3: Desigualdad de Cauchy-Schwarz

Sean v y w vectores en R^n, entonces:

$$|v \cdot w| \leq |v|\,|w|$$

Demostración

Para escalares r y s cualesquiera se tiene:

$$|rv + sw|^2 \geq 0 \quad (5.1)$$

$$|rv + sw|^2 = (rv+sw)(rv+sw) = r^2(vv) + 2rs(vw) + s^2(ww) \quad (5.2)$$

Haciendo $r=ww$ y $s = -(vw)$ se obtiene de las ecuaciones (5.1) y (5.2)

$$(ww)^2(vv) - 2(ww)(vw)(vw) + (vw)^2(ww) = (ww)^2(vv) - (ww)(vw)^2 \geq 0$$

$$(ww)\left[(vv)(ww)-(vw)^2\right] \geq 0$$

Si $(ww) = 0$ entonces $w=0$ y se tiene la igualdad con cero. Si $(ww) \neq 0$, entonces: $ww>0$ y se tiene:

$$(vv)(ww)-(vw)^2 \geq 0$$

$$(vv)(ww) \geq (vw)^2$$

$$|v|^2\,|w|^2 \geq (vw)^2$$

Tomando raíces cuadradas se tiene:

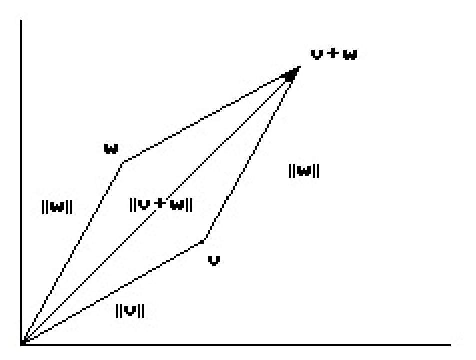
$$|v|\,|w| \geq |v\,w|$$

Otra desigualdad importante que es derivada de la desigualdad de Cauchy–Schwarz es la desigualdad del triángulo: dados v y w, w en R^n se tiene:

$$|v + w| \leq |v| + |w|$$

La figura 5.1 muestra el origen del nombre de desigualdad triangular

Figura 5.1 Desigualdad triangular



Demostración
Usando las propiedades del producto punto, así como la desigualdad de Cauchy– Schwarz tenemos:

$$|v+w|^2 = (v+w)(v+w) = v\cdot v + 2v\cdot w + w\cdot w$$
$$\leq v\cdot v + 2|v|\,|w| + ww= |v|^2 + 2|v|\,|w| + |w|^2$$
$$= \left(|v| + |w|\right)^2$$

Y tomando raíces cuadradas se tiene:

$$|v+w| \leq |v| + |w|$$

Ya se han visto algunas propiedades de los vectores en R^n, ahora procederemos a generalizarlas y definir un espacio vectorial.

Definición 5.3
Un espacio vectorial real es un conjunto no vacío V de objetos llamados vectores junto con una regla para sumar dos vectores cualesquiera v y w para producir un vector $v + w \in V$ y una regla para multiplicar cualquier vector $v \in V$ por cualquier escalar $r \in R$ a fin de producir un vector $r\, v \in V$ llamado múltiplo escalar. Además, para todos los vectores u, v y w en V y para todos los escalares r y s, se deben cumplir los siguientes axiomas:

Propiedades de la suma

1. Si u y v son vectores en V entonces $u + v \in V$ (cerrado bajo la suma)
2. $u+v= v+u$ (ley conmutativa)
3. $u + (v+w)= (u+v) + w$ (ley asociativa)
4. Existe un vector $0 \in V$ llamado vector cero
   $0+u= u+0= u$ para toda $u \in V$ (naturaleza del vector cero)
5. Para toda $u \in V$ existe $-u \in V$ tal que
   $u +(-u)= (-u)+u= 0$ (–u como inverso aditivo de u)

Propiedades que incluyen el producto por un escalar

6. Sí $r \in R$ y $u \in V$, entonces $ru \in V$ (cerrado bajo el producto por un escalar)
7. $r(u+v)= ru+rv$ (ley distributiva)

8. $(r + s)u= ru + su$ (ley distributiva)
9. $r(su)= (rs)u$ (ley asociativa)
10. $1v= v$ (preservación de escala)

Ejemplo 5.3

El conjunto $V=R^n$ con las operaciones ordinarias de adición y multiplicación por escalares es un espacio vectorial.

Ejemplo 5.4

Sea V cualquier plano que pasa por el origen en R^3. Los puntos de V forman un espacio vectorial bajo las operaciones ordinarias para vectores en R^3. Por el ejemplo anterior se sabe que R^3es un espacio vectorial bajo esas operaciones. Los axiomas (2) (3) (7) (8) (9) y (10) se cumplen para todos los puntos de R^3 y por lo tanto para todos los puntos del plano V. Entonces solo es necesario demostrar que se cumplen los axiomas (1) (4) (5) y (6).

Solución

Dado que el plano V pasa por el origen tiene una ecuación de la forma:

$$ax+by+cz= 0 \tag{5.3}$$

Por lo tanto, si $v = (v_1, v_2, v_3)$ y $w= (w_1, w_2, w_3)$ son puntos en V entonces:

$$av_1+bv_2+cv_3= 0$$
$$aw_1+bw_2+cw_3= 0$$

Sumando estas ecuaciones se tiene:

$$a(v_1+w_1) + b(v_2+w_2) + c(v_3+w_3)= 0$$

Esta igualdad indica que las coordenadas del punto $v +w = (v_1+w_1, v_2+w_2, v_3+w_3)$ satisfacen (5.3) y por lo tanto $v+w\in V$, lo que muestra que el axioma 1 se cumple.

Si multiplicamos la ecuación $av_1+bv_2+cv_3= 0$ por -1 resulta:

$$a(-v_1) + b(-v_2) + c(-v_3)= 0$$

Por lo tanto $-v \in V$ y se cumple el axioma 5. La demostración de los axiomas 4 y 6 se dejan como ejercicio.

Ejemplo 5.5

Muestre que el conjunto P de todos los polinomios con coeficientes en R es un espacio vectorial, usando como suma de vectores y producto por un escalar la suma usual de polinomios y el producto usual de un polinomio por un escalar.

Solución

Sean p y q polinomios:

$$p= a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$
$$q= b_0 + b_1x + b_2x^2 + \cdots + b_mx^m$$

Si $m \geq n$, recordemos que la suma de p y q está dada por:

$$p+q = (a_0+b_0) + (a_1+b_1)x + \cdots + (a_n+b_n)x^n + \ b_{n+1}x_{n+1} + \cdots + b_mx^m$$

Se hace una definición similar si $m<n$. El producto de p por un escalar r está dado por:

$$rp= ra_0 + ra_1x + ra_2x^2 + \cdots + ra_nx^n$$

Se tienen el polinomio cero y $-p$ el inverso aditivo, entonces se cumplen los 8 axiomas restantes.

Ejemplo 5.6

Demuestre que el conjunto M de todas las matrices de $m \times n$ es un espacio vectorial, usando como suma de vectores y multiplicación por un escalar la suma y producto por un escalar usual para las matrices.

Solución

La suma de matrices de $m \times n$ y el producto de una matriz de $m \times n$ por un escalar produce de nuevo una matriz de $m \times n$. Así M es cerrado bajo la suma de vectores y el producto por un escalar. Tomamos como vector cero en M la matriz cero usual, cuyas entradas son todas iguales a cero. Para cualquier matriz A en M consideremos $-A$ como la matriz $(-1)A$.

Ejemplo 5.7
Este ejemplo muestra que no necesariamente la suma vectorial ha de estar relacionada con la suma ordinaria, y que el vector cero no necesariamente es el número real cero.

Sea $V = \{ x \mid x \in R,\ x > 0 \}$ la suma y producto escalar se definen como:

$x \oplus y = x\, y$
suma vectorial = producto ordinario

Y para $r \in R$, $x \in V$, se define:

$r \otimes x = x^r$(potencia ordinaria)

Muestre que V con estas operaciones es un espacio vectorial.

Solución
Como los elementos de V son $x \in$R entonces se tiene:

$2 \oplus 3 = 2 \cdot 3 = 6$
$2 \otimes 3 = 3^2 = 9$
$-2 \otimes 3 = 3^{-2} = 1/9$

Entonces tenemos que demostrar que se satisfacen los axiomas:

1. Cerradura para la suma.
   Sean $x \in V$, $y \in V$ entonces $x \oplus y = x\, y$ como $x, y \in R^+$ y como el producto de dos reales positivos es otro real positivo $x\, y \in V$ entonces $x \oplus y \in V$.

2. Cerradura para la multiplicación por un escalar.
   Sean $r \in R$ y $x \in V$ $r \otimes x = x^r$ como $x \in R^+$ al elevarlo a cualquier potencia real se obtiene un número real positivo. En consecuencia, $r \otimes x \in V$.

3. Conmutatividad de la suma:
   $x \oplus y = xy = yx = y \oplus x$
   por lo tanto: $x \oplus y = y \oplus x$

4. Asociatividad:
   $(x \oplus y) \oplus z = (xy) \oplus z = (xy)z = x(yz) = x \oplus (yz) = x \oplus (y \oplus z)$

5. Idéntico aditivo:
   $x \oplus \theta = x$ por definición $x \oplus \theta = x\theta$ de donde $\theta = 1$, es decir el real positivo 1 es el idéntico aditivo o cero para este espacio.
   $x \oplus 1 = x$

6. Inverso aditivo.
   Para toda $x \in V$, $-x$ debe satisfacer $x \oplus (-x) = \theta$ pero $-x$ no es "menos x" es el símbolo para el inverso aditivo y θ no es el cero, sino 1 entonces:
   $x \oplus (-x) = \theta$

   Es equivalente a$x(-x) = 1$

   De donde:
   $-x = 1/x$ como $x > 0$ entonces $1/x > 0$ por lo tanto $-x \in V$

7. Ley asociativa para la multiplicación escalar.
   Sea $r \in R, s \in R, x \in V$ entonces:
   $(rs) \otimes x = x^{rs}$ y $r \otimes (s \otimes x) = r \otimes (x^s) = (x^s)^r$

   Que por las leyes de los exponentes:
   $(x^s)^r = x^{sr} = x^{rs}$ en consecuencia,
   $(rs) \otimes x = r \otimes (s \otimes x)$

8 y 9. Leyes distributivas. Sean $r \in R, s \in R, x \in V, y \in V$ entonces:
   $(r+s) \otimes x = x^{r+s} = x^r x^s = x^r \oplus x^s = (r \otimes x) \oplus (s \otimes x)$
   $r \otimes (xy) = r \otimes (xy) = (xy)^r = x^r y^r = x^r + y^r = (r \otimes x) \oplus (r \otimes y)$

10. Sea $x \in V$, entonces
   $1 \otimes x = x^1 = x$

5.3. SUBESPACIOS

Hemos visto que R^n es un espacio vectorial en el que la suma de vectores consiste en sumar las componentes correspondientes de tales vectores y el producto por un escalar se obtiene multiplicando cada componente por el escalar. Cualquier subconjunto de R^n que con estas mismas operaciones sea también un espacio vectorial, se llama subespacio de R^n.

Por ejemplo, el plano x_1x_2 de R^3 formado por todos los vectores que tienen cero como tercera componente (como se puede ver en la figura 5.2), es un subespacio de R^3.

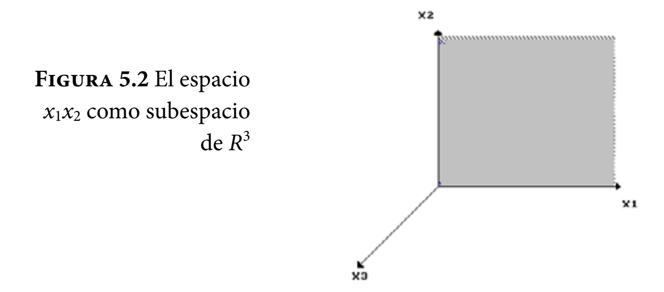


Figura 5.2 El espacio x_1x_2 como subespacio de R^3

Por su parte, el subconjunto $\{(m, n,p) \mid m, n, p \in Z\}$ de R^3 que consta de todos los vectores de R^3 con componentes enteras no es un subespacio, pues, aunque este subconjunto es cerrado bajo la suma de vectores, no es cerrado bajo el producto por un escalar, por ejemplo:

0.5 (1, 2, 5) = (0.5, 1, 2.5) no está en el subconjunto

Definición 5.4

Un subconjunto W de un espacio vectorial V es un subespacio de V si W es por si solo un espacio vectorial bajo las operaciones de suma de vectores y de producto por un escalar definidas en V.

Esto significa que el conjunto W debe ser cerrado bajo la suma y el producto por un escalar. Las propiedades requeridas por un espacio vectorial se cumplen para W pues son válidas para todo V.

Teorema 5.4: Criterio para un subespacio

Un subconjunto no vacío W de un espacio vectorial V es un subespacio de V si y solo si satisface las siguientes condiciones:

a) Si u y v son vectores de W entonces $u + v$ está en W

b) Si r es cualquier escalar de R y v está en W, entonces $r\,v$ está en W

Las condiciones a) y b) indican que W es cerrado bajo la suma y bajo el producto por un escalar.

Demostración

⇨)

Si W es un subespacio de V, entonces satisface todos los axiomas de espacio vectorial, en particular, cumple con los axiomas 1 y 6 pero estos axiomas son las condiciones a y b.

⇦)

Suponemos que se cumplen a y b. Dado que estas condiciones son los axiomas 1 y 6 de espacio vectorial, solamente resta demostrar que W satisface los otros 8 axiomas. Los axiomas 2, 3, 7, 8, 9, 10 se satisfacen automáticamente ya que todos los vectores en V los satisfacen. Para completar la demostración solo falta verificar que W satisface los axiomas 4 y 5.

4) Sea $u \in W$, por la condición b, $ru \in W$ para toda $r \in R$. Haciendo $r = -1$ se tiene:

$(-1)\,u \in W$ de donde

$(-1)u + u \in W$ por a)

Y esto implica que:

$(-1)u + u = 0$ por lo que $0 \in W$

5) Usando 4 consideramos $r = -1$ y $(-1)u = -u \in W$

Todo espacio vectorial V tiene al menos 2 subespacios. V es un subespacio de sí mismo y el conjunto $\{0\}$ que consiste únicamente del vector cero en V, este subespacio recibe el nombre de subespacio trivial. Los subespacios que no son ni $\{0\}$ ni V se conocen como **subespacios propios**.

El teorema 5.3 dice que para probar que *W* es un subespacio de *V* basta verificar que se cumplen las condiciones *a* y *b*.

Algunos ejemplos de subespacios son:

Ejemplo 5.8

Para todo espacio vectorial *V*, el subconjunto {0} que contiene solamente el vector cero, es un subespacio, puesto que:

$0 + 0 = 0$

$r\,0 = 0$ para toda $r \in R$

Y se le conoce como subespacio trivial.

Ejemplo 5.9

El conjunto $M_{22}{}^{0}$ de todas las matrices de 2×2 que tiene ceros en la diagonal principal es un subespacio del espacio vectorial M_{22} de todas las matrices de orden 2×2.

Solución

Sean *A* y *B* dos matrices en $M_{22}{}^{0}$ como sigue:

$$A=\begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix} \qquad B=\begin{bmatrix} 0 & b_{12} \\ b_{21} & 0 \end{bmatrix}$$

Entonces $r\,A$ y $A + B$ son iguales a:

$$rA=\begin{bmatrix} 0 & ra_{12} \\ ra_{21} & 0 \end{bmatrix} \qquad A+B=\begin{bmatrix} 0 & a_{12}+b_{12} \\ a_{21}+b_{21} & 0 \end{bmatrix}$$

Que como se ve tienen ceros en la diagonal principal, entonces pertenecen a $M_{22}{}^{0}$. Por lo tanto, $M_{22}{}^{0}$ es un subespacio de M_{22}.

Ejemplo 5.10

Sea $Ax=0$ un sistema homogéneo de m ecuaciones lineales con n incógnitas y sea *W* el conjunto de todos los vectores solución del sistema y sean *u* y *v* dos vectores en *W*. Se demostrará que *W* es un subespacio de R^n.

Solución

Se dice que un vector $v \in R^n$ es un vector solución del sistema si:

$x_1 = v_1, x_2 = v_2, \cdots, x_n = v_n$ es una solución del sistema.

Para mostrar que es cerrado bajo la suma y bajo el producto por un escalar, es necesario demostrar que $u+v$ y ru son vectores solución.

Dado que u y v son vectores solución se tiene que:

$Au = 0 \quad \text{y} \quad Av = 0$

Y de aquí:

$A(u+v) = Au+Av = 0 + 0 = 0$

Entonces: $u+v$ satisface $Ax = 0$ lo que implica $u + v \in W$, además:

$A(ru) = r(Au) = r0 = 0$

Se puede concluir que $r\,u$ satisface $Ax = 0$ entonces $ru \in W$, por lo tanto, W es un subespacio de R^n.

Ejemplo 5.11

Sea $H = \{(x, y) \mid y = 2x,\ x, y \in R\}$ esto es, H consiste en los vectores en R^2 que están sobre una recta que pasa por el origen. Verificaremos que H es un subespacio de R^2.

Solución

Sean:

$a = (x_1, y_1) \in H$ entonces $y_1 = 2x_1$

$b = (x_2, y_2) \in H$ entonces $y_2 = 2x_2$

$a+b = (x_1+x_2, y_1+y_2) = (x_1+x_2, 2x_1+2x_2) = (x_1+x_2, 2(x_1+x_2))$

De aquí se concluye que $a+b \in H$:

$ra = r(x_1, y_1) = (rx_1, ry_1) = (rx_1, r2x_1) = (rx_1, 2(rx_1))$

Entonces se concluye que $ra \in H$ y por lo tanto H es un subespacio de R^2.

Ejemplo 5.12

Considere todos los vectores en R^2 cuyas componentes sean positivas o cero. Si el espacio vectorial V es R^2 con las operaciones usuales de suma y producto para vectores entonces este subconjunto es el primer cuadrante, las coordenadas satisfacen $x \geq 0$, $y \geq 0$.

Este subconjunto sin embargo no es un subespacio ya que aunque contenga al cero y sea cerrado bajo la suma, no es cerrado bajo el producto ya que si $r = -1$ se tiene $r(x,y) = -1(x,y) = (-x,-y)$ que se encuentra en el tercer cuadrante. Si se incluye el tercer cuadrante además del primero, no es cerrado bajo la suma ya que:

$(1, 2) + (-2, -1) = (-1, 1)$ que está en el segundo cuadrante.

Por lo tanto, el menor subespacio que contiene al primer cuadrante es todo el espacio R^2.

Nota: No todo espacio vectorial tiene subespacios propios.

Ejemplo 5.13

Sea W un subespacio de R. Si $W \neq \{0\}$ entonces W contiene un número real $a \neq 0$ por el axioma 6 existe $1/\alpha$tal que:

$(1/\alpha)\ \alpha = 1 \in W$

y $1\alpha = \alpha \in W$ para toda $\alpha \in R$

Así si W no es el subespacio trivial entonces $W = R$, esto es, R no tiene subespacios propios.

Se puede pensar que cualquier recta o plano de R^3 es un subespacio, sin embargo, para que esto sea cierto, este subespacio de R^3 debe contener al cero. Geométricamente significa que esta recta o plano debe pasar por el origen.

5.4. COMBINACIONES LINEALES

Como ya se decía, un plano o una recta en R^3 es un subespacio solo si pasa por el origen. Un plano que pasa por el origen está totalmente determinado por dos vectores cualesquiera v_1 y v_2 distintos de cero y no paralelos que están en él, como se ve en la figura 5.3.

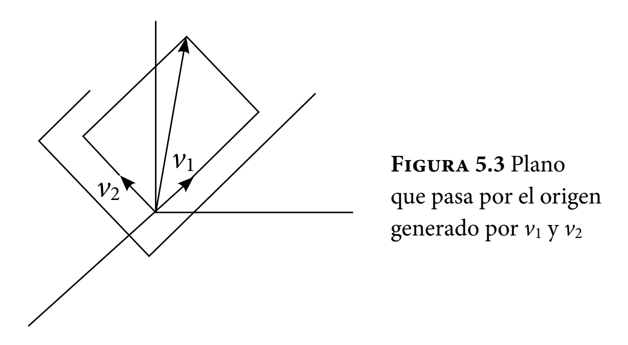


Figura 5.3 Plano que pasa por el origen generado por v_1 y v_2

La figura muestra que cualquier vector de este plano tiene la forma:

$$r_1 v_1 + r_2 v_2$$

Esta expresión es una combinación lineal de los vectores v_1 y v_2.

Definición 5.5

Se dice que un vector w es una combinación lineal de los vectores $v_1, v_2, \cdots, v_n$ si es posible expresarlo en la forma:

$$w = r_1 v_1 + r_2 v_2 + \cdots + r_n v_n$$

Donde $r_1, r_2, \cdots, r_n$ son escalares.

Ejemplo 5.14

Sean los vectores $u=(1, 2, -1)$ y $v=(6, 4, 2)$ en R^3. Demuestre que $w=(9, 2, 7)$ es una combinación lineal de u y v.

Solución

Para que w sea una combinación lineal de u y v, deben existir escalares r_1 y r_2 tales que:

$$w = r_1\, u + r_2 v$$

Es decir:

$$(9, 2, 7) = r_1\,(1, 2, -1) + r_2\,(6, 4, 2)$$

Que equivale a:

$(9, 2, 7) = (r_1 + 6r_2, 2r_1 + 4r_2, -r_1 + 2r_2)$

Igualando las componentes correspondientes resulta:

$r_1 + 6r_2 = 9$
$2r_1 + 4r_2 = 2$
$-r_1 + 2r_2 = 7$

Resolviendo este sistema se obtiene $r_1 = -3$ y $r_2 = 2$ de donde

$w = -3u + 2v$

En general definimos:

Definición 5.6
Sea $\{v_1, v_2, \cdots, v_n\}$ un conjunto de vectores en el espacio vectorial V. El espacio generado por $\{v_1, v_2, \cdots, v_n\}$ es el conjunto de todas las combinaciones posibles de $v_1, v_2, \cdots, v_n$ y se denota:

$$gen(v_1, v_2, \cdots, v_n) = \{r_1 v_1 + r_2 v_2 + \cdots + r_n v_n \mid r_i \in R\ i=1,2,..,n\}$$

Ejemplo 5.15
Para el ejemplo del plano en R^3 tendríamos que el conjunto de todas las combinaciones lineales de v_1 y v_2 sería:

$$gen\ (v_1, v_2) = \{r_1 v_1 + r_2 v_2 \mid r_1, r_2 \in R\}$$

Y geométricamente se puede ver en la figura 5.4. Además, $gen\ (v_1, v_2, \cdots, v_n)$ V es un subespacio de V.

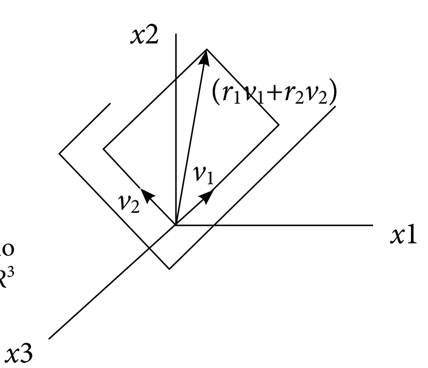


Figura 5.4 Subespacio generado por v_1 y v_2 en R^3

CONTENIDO

Teorema 5.5

Si $S=\{v_1, v_2, \cdots, v_n\} \subset$ siendo V un espacio vectorial, entonces *gen S* es un subespacio de V.

Demostración

Sean $x, y \in$ *gen S*, entonces x, y son combinaciones lineales:

$$x = r_1 v_1 + \cdots + r_n v_n$$
$$y = s_1 v_1 + \cdots + s_n v_n$$

Entonces:

$$x+y= (r_1+s_1)v_1 + \cdots + (r_n+s_n) v_n$$

y

$$kx= (kr_1)v_1 + \cdots + (kr_n) v_n$$

que también se encuentra en *gen S*. Por lo tanto, *gen S* es un subespacio de V. Se le llama subespacio de V generado por S.

Ejemplo 5.16

Determine si $v_1= (1, 1, 2)$, $v_2= (1, 0, 1)$ y $v_3= (2, 1, 3)$ generan a R^3.

Solución

Es necesario determinar si un vector arbitrario $b= (b_1, b_2, b_3)$ en R^3 se puede expresar como una combinación lineal de v_1, v_2 y v_3 esto es:

$$b= r_1v_1 + r_2v_2 + r_3v_3$$

En términos de sus componentes resulta:

$$(b_1, b_2, b_3) = (r_1 + r_2 + 2r_3, r_1 + r_3, 2r_1 + r_2 + 3r_3)$$

O bien:

$$\begin{aligned} r_1 + r_2 + 2r_3 &= b_1 \\ r_1 + \quad + r_3 &= b_2 \\ 2r_1 + r_2 + 3r_3 &= b_3 \end{aligned}$$

Entonces el problema se reduce a determinar si este sistema es consistente para todos los valores de b_1, b_2 y b_3. Esto es equivalente a que la matriz de los coeficientes sea invertible.

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

O lo que es lo mismo $det\ A \neq 0$, como $det\ A=0$, A no es invertible y por lo tanto v_1, v_2 y v_3 no generan a R^3.

Ejemplo 5.17

Describa geométricamente el subespacio de R^4 generado por el vector $v=(2, 3, -1, 4)$.

Solución

El subespacio $gen(v)$ es una recta de R^4 que contiene al origen, cada punto de esta recta es un múltiplo escalar rv de v, esto es, cada punto tiene la forma (x_1, x_2, x_3, x_4) donde:

$$x_1 = 2r,\ x_2 = 3r,\ x_3 = -r \text{ y } x_4 = 4r \text{ para algún escalar } r.$$

Nota: Que un determinado vector u se pueda expresar como una combinación lineal de otro u otros vectores v_i se puede ver también, como el hecho de que ese vector u pertenezca al espacio generado por los vectores v_i.

Ejemplo 5.18

¿Es posible escribir $(3,-1, 4)$ como combinación lineal de $(0, 1, 1)$, $(1,-1, 0)$ *y* $(3, -5,-2)$? O lo que es lo mismo ¿está $(3,-1, 4)$ en gen $\{(1,-1, 0), (0, 1, 1), (3,-5,-2)\}$?

Solución

En cualquier caso, se tiene:

$$(3, -1, 4) = r_1\,(1, -1, 0) + r_2\,(0, 1, 1) + r_3\,(3, -5, -2)$$

Equivalentemente:

$$\begin{aligned} 3 &= r_1 + \quad + 3r_3 \\ -1 &= -r_1 + r_2 - 5r_3 \\ 4 &= \qquad r_2 - 2r_3 \end{aligned}$$

Escribiendo en forma matricial y efectuando eliminación gaussiana se tiene:

$$A = \left[\begin{array}{ccc|c} 1 & 0 & 3 & 3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

Entonces no existe solución y (3, –1, 4) no se puede escribir como combinación lineal de los vectores dados.

5.5. INDEPENDENCIA LINEAL

Hemos visto que un conjunto de vectores $gen(S)$ genera un espacio vectorial V si $S \subset$ y cada vector en V es una combinación lineal de $v_1, v_2, \cdots, v_r$. Los conjuntos de generadores son muy útiles, ya que en muchos casos es posible estudiar un espacio vectorial V si primero se trabaja con los vectores que pertenecen a algún conjunto de generadores S y después se aplican los resultados a la totalidad de V. El problema de encontrar los conjuntos más pequeños de generadores para un espacio vectorial depende del concepto de independencia lineal.

Ejemplo 5.19
Determine si $gen\left\{(1, -1, 0), (0, 1, 1), (3, -5, -2)\right\}$ genera todo R^3.

Solución
Sea (x_1, x_2, x_3) un vector cualquiera en R^3, entonces queremos saber si es posible escribir:

$$(x_1, x_2, x_3) = r_1(1, -1, 0) + r_2(0, 1, 1) + r_3(3, -5, -2)$$

Que es equivalente a:

$$A = \left[\begin{array}{rrr|l} 1 & 0 & 3 & x_1 \\ -1 & 1 & -5 & x_2 \\ 0 & 1 & -2 & x_3 \end{array}\right]$$

La cual se reduce a:

$$A = \left[\begin{array}{rrr|lll} 1 & 0 & 3 & x_1 & & \\ 0 & 1 & -2 & x_1 & +x_2 & \\ 0 & 0 & 0 & x_3 & -x_2 & -x_1 \end{array}\right]$$

Hay una solución solo si $x_3-x_2-x_1=0$, pero esto impone una restricción a (x_1, x_2, x_3) y por lo tanto la primera de las ecuaciones no se puede resolver para un vector arbitrario (x_1, x_2, x_3), esto significa que este conjunto de vectores no genera R^3 lo que no implica que no se pueda generar otro espacio con este conjunto, entonces se puede describir gen(S) donde:

$$S = \{(1, -1, 0), (0, 1, 1), (3, -5, -2)\}$$

Suponga que (x_1, x_2, x_3) está en *gen*(S) entonces la ecuación:

$$(x_1, x_2, x_3) = r_1(1, -1, 0) + r_2(0, 1, 1) + r_3(3, -5, -2)$$

Debe tener solución, pero se tiene que $x_3 - x_2 - x_1 = 0$ por lo cual:

$$gen(\mathrm{S}) = \{(x_1, x_2, x_3) | \, x_3 = x_1 + x_2\}$$

Es decir, el espacio generado por S consta de todos los vectores cuya tercera componente es la suma de las dos primeras, así por ejemplo (1, 3, 5) no está en *gen*(S) y $(1, 3, 4) \in gen(S)$.

Por ejemplo, R^2 está generado por $S = \{(1, 0), (0, 1)\}$ pero también puede generarse por el conjunto mayor $S' = \{(1, 0), (0, 1), (1, 1)\}$. R^2 y funciones sobre R^2 se pueden analizar a través de los conjuntos generadores, así entonces por economía, habría que encontrar los conjuntos generadores más pequeños. Para lograr esto se requiere del concepto de **independencia lineal**.

Definición 5.7
Un conjunto $\{v_1, v_2, \cdots, v_n\}$ de vectores en un espacio vectorial V se llama linealmente independiente si:

$$r_1 v_1 + r_2 v_2 + \cdots + r_n v_n = 0$$

Tiene como única solución $r_1 = r_2 = \cdots = r_n = 0$, es decir la solución trivial.

Definición 5.8
Un conjunto $\{v_1, v_2, \cdots, v_n\}$ de vectores en un espacio vectorial V se llama linealmente dependiente si:

$$r_1 v_1 + r_2 v_2 + \cdots + r_n v_n = 0$$

Tiene una solución no trivial, es decir, existen escalares $r_1, r_2, \cdots, r_n$ no todos cero tales que el sistema descrito tiene solución.

Ejemplo 5.20
En cada uno de los espacios vectoriales especificados determine si los vectores listados son linealmente independientes o dependientes.

a) $V = R^2$ con $S = \{x_1 = (2, 1), x_2 = (6, 3), x_3 = (0, 1)\}$

Geométricamente se tiene:

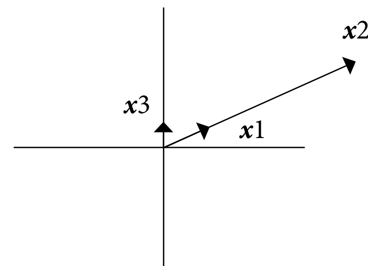


Figura 5.5 Espacio generado con tres vectores en R^2

Note que

$$-3x_1 + x_2 + 0x_3 = 0$$

De donde x_1, x_2, x_3 son linealmente dependientes.

Otra forma sería:

$r_1 x_1 + r_2 x_2 + r_3 x_3 = 0$

$$r_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + r_2 \begin{bmatrix} 6 \\ 3 \end{bmatrix} + r_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Y en forma matricial:

$$\begin{bmatrix} 2 & 6 & 0 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

De donde:

$$\begin{bmatrix} 2 & 6 & 0 \\ 1 & 3 & 1 \end{bmatrix} \approx \begin{bmatrix} 2 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Lo que implica $r_3 = 0$, $r_2 = t$, $r_1 = -3t \quad t \in R$.

Por lo tanto: x_1, x_2, x_3 son linealmente dependientes.

b) Si $V = R^3$ con:

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \quad x_3 = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

Entonces:

$$r_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + r_2 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + r_3 \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

O bien:

$$A = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 1 & 4 & 0 \end{array}\right]$$

Como se tiene una matriz triangular inferior es invertible y además la solución es la trivial.

$r_1 = r_2 = r_3 = 0$

Por lo cual, x_1, x_2, y x_3 son linealmente independientes.

c) Si $V=M_{2\times 2}$ con:

$$M_1=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad M_2=\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \quad M_3=\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Entonces escribimos:

$$r_1M_1 + r_2M_2 + r_3M_3 = 0$$

Esto es:

$$r_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + r_2 = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} + r_3 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

O bien:

$$\begin{bmatrix} r_1+4r_2 & r_3 \\ r_3 & r_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

De aquí:

$$\begin{aligned} r_1 + 4r_2 &= 0 \\ r_3 &= 0 \end{aligned}$$

Lo que implica que $r_1=-4r_2$ si $r_2=t$, entonces: $r_1=-4t$, $t \in R$

Por lo tanto, M_1, M_2, M_3 son linealmente dependientes.

Ejemplo 5.21

Determine si el conjunto formado por los vectores $v_1 = (1, -2, 3)$, $v_2 = (5, 6, -1)$ y $v_3 = (3, 2, 1)$ es linealmente dependiente o independiente.

Solución

En términos de los componentes, la ecuación vectorial

$$r_1 v_1 + r_2 v_2 + r_3 v_3 = 0$$

se transforma en:

$$r_1 (1, -2, 3) + r_2 (5, 6, -1) + r_3 (3, 2, 1) = (0, 0, 0)$$

O el sistema:

$$\begin{aligned} r_1 + 5r_2 + 3r_3 &= 0 \\ -2r_1 + 6r_2 + 2r_3 &= 0 \\ 3r_1 - r_2 + r_3 &= 0 \end{aligned}$$

Resolviendo este sistema se tiene:

$$r_1 = -\tfrac{1}{2}\,t \qquad r_2 = -\tfrac{1}{2}\,t \qquad r_3 = t$$

Por lo cual, el sistema tiene soluciones no triviales y en consecuencia v_1, v_2 y v_3 son linealmente dependientes.

Nota: El término *linealmente dependiente* sugiere que los vectores *dependen* unos de otros de alguna manera. Para ver esta dependencia, suponga que $v= \{v_1, v_2, \cdots, vr\}$ es un conjunto linealmente dependiente, por lo tanto, la ecuación:

$$r_1 v_1 + r_2 v_2 + \cdots + r_n v_n = 0$$

Tiene otra solución además de $r_1=r_2 = \cdots = r_r = 0$. Supongamos que $r_1 \neq 0$ entonces multiplicando esta ecuación por $1/r_1$ y resolviendo para v_1 se obtiene:

$$v_1 = \left(-\frac{r_2}{r_1}\right) v_2 + \cdots\cdots + \left(-\frac{r_r}{r_2}\right) v_r$$

Por lo que v_1 se puede expresar como una combinación lineal de los vectores restantes $v_2, \cdots, v_r$. Por lo que podemos concluir que: un conjunto de vectores es linealmente dependiente si y solo si al menos uno de los vectores es una combinación lineal de los restantes.

Ejemplo 5.22

Dos vectores v_1 y v_2 forman un conjunto linealmente dependiente, si y solo si uno de los vectores es un múltiplo escalar del otro. Para ver esto, suponga que $S=\{v_1, v_2\}$ es linealmente dependiente. Dado que la ecuación vectorial:

$$r_1 v_1 + r_2 v_2 = 0$$

Tiene una solución además de la trivial $r_1=r_2=0$, esta ecuación se puede expresar como:

$$v_1 = -\left(\frac{r_2}{r_1}\right) v_2 \quad \text{o} \quad v_2 = -\left(\frac{r_1}{r_1}\right) v_1$$

Estas ecuaciones indican que v_1 es múltiplo escalar de v_2 y viceversa. Por consecuencia, dos vectores en R^2 o en R^3 son linealmente dependientes, si y solo si pertenecen a la misma recta que pasa por el origen.

Para R^3 decimos que si v_1, v_2 y v_3 son tres vectores en R^3. Entonces, el conjunto $S = \{v_1, v_2, v_3\}$ es linealmente dependiente si y solo si los tres vectores pertenecen al mismo plano que pasa por el origen.

Teorema 5.6

Sea $V=\{v_1, v_2, \cdots, v_r\}$ un conjunto de vectores en R^n. Si $r > n$, entonces V es linealmente dependiente.

Demostración

Suponga que:

$$\begin{aligned} v_1 &= (v_{11}, v_{12}, \cdots, v_{1n}) \\ v_2 &= (v_{21}, v_{22}, \cdots, v_{2n}) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ v_r &= (v_{r1}, v_{r2}, \cdots, v_{rn}) \end{aligned}$$

Considere la ecuación:

$$r_1 v_1 + r_2 v_2 + \cdots + r_r v_r = 0$$

Expresando ambos lados de la ecuación en términos de sus componentes se tiene que:

$$v_{11}\, r_1 + v_{21}\, r_2 + \cdots + v_{r1} r_r = 0$$
$$v_{12}\, r_1 + v_{22}\, r_2 + \cdots + v_{r2} r_r = 0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$v_{1n}\, r_1 + v_{2n}\, r_2 + \cdots + v_{rn}\, r_r = 0$$

Este es un sistema homogéneo de n ecuaciones con r incógnitas. Dado que $r>n$ entonces por el resultado que dice que, si un sistema homogéneo de ecuaciones lineales tiene más incógnitas que ecuaciones, siempre tiene un número infinito de soluciones, lo que implica que el sistema tiene soluciones no triviales y por lo tanto S es un conjunto linealmente dependiente.

5.6. BASES Y DIMENSIÓN

El conjunto de vectores $S = \{(1, 1), (1, -1)\}$ genera a R^2. Es decir, cualquier vector de R^2es una combinación lineal de (1, 1), (1,–1). El conjunto de vectores:

$$T = \{(1, 1), (1, -1), (1, 0)\}$$

También genera a R^2. Los conjuntos S y T difieren en que S es linealmente independiente, mientras que T es linealmente dependiente. Esto hace que haya una diferencia al expresar un vector como una combinación lineal de los vectores de cada conjunto. Por ejemplo, al expresar (2, 4) en términos de los vectores en S, se tiene como única posibilidad:

$$(2, 4) = 3(1, 1) - 1(1, -1)$$

Sin embargo, en términos de los vectores de T, hay varias posibilidades:

$$(2, 4) = 3(1, 1) - 1(1, -1) + 0(1, 0)$$
$$(2, 4) = 0(1, 1) - 4(1, -1) + 6(1, 0)$$
$$(2, 4) = 4(1, 1) + 0(1, -1) - 2(1, 0)$$

En general:

$$(2, 4) = (k+4)(1, 1) + k(1, -1) + (-2-2k)(1, 0)$$

La conclusión es: si un conjunto S de vectores genera V y S es linealmente dependiente, entonces la representación de un vector x en términos de los vectores en S no es única. Si se desea unicidad, el conjunto generador también deberá ser linealmente independiente. Un conjunto con esas características recibe el nombre de **base** para V. Las bases se emplean en teoría de códigos entre otras áreas.

Definición 5.9
Si V es un espacio vectorial y $S = \{v_1, v_2, \cdots, v_r\}$ es un conjunto finito de vectores en V, entonces se dice que S es una base de V si:
a) S es linealmente independiente
b) S genera a V

Ejemplo 5.23
Sean $v_1 = (1, 2)$ y $v_2 = (3, -1)$. Demuestre que el conjunto $S = \{v_1, v_2\}$ es una base de R^2.

Solución
Se tiene que demostrar que S es linealmente independiente y que genera a R^2.

Para demostrar que S genera a R^2, se debe mostrar que un vector arbitrario $b=(b_1, b_2)$ se puede expresar como una combinación lineal de v_1 y v_2, esto es:

$$r_1 (1, 2) + r_2 (3, -1) = (b_1, b_2) \qquad (5.4)$$

Tiene una solución única para cualquier $b=(b_1, b_2)$.

Y para demostrar que v_1 y v_2 son linealmente independientes se debe demostrar que:

$$r_1 (1, 2) + r_2 (3, -1) = (0, 0) \qquad (5.5)$$

Tiene solamente la solución $r_1=r_2=0$. Las ecuaciones (5.4) y (5.5) escritas en términos de sus componentes son iguales a:

$$\begin{aligned} r_1+3r_2&=b_1 \\ 2r_1-r_2&=b_2 \end{aligned} \quad \text{y} \quad \begin{aligned} r_1+3r_2&=0 \\ 2r_1-r_2&=0 \end{aligned}$$

En forma matricial se tiene:

$$\left[\begin{array}{cc|c} 1 & 3 & b_1 \\ 2 & -1 & b_2 \end{array}\right] \left[\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & -1 & 0 \end{array}\right]$$

Lo que se puede resolver escribiendo:

$$\left[\begin{array}{cc|c|c} 1 & 3 & b_1 & 0 \\ 2 & -1 & b_2 & 0 \end{array}\right]$$

Usando eliminación gaussiana se tiene:

La independencia lineal se comprueba con esta parte

$$\left[\begin{array}{cc|c|c} 1 & 3 & b_1 & 0 \\ 0 & -1 & \frac{-b_2 + 2b_1}{7} & 0 \end{array}\right]$$

El espacio generado se comprueba con esta parte

Entonces, para demostrar que S genera a R^2 tenemos que:

$$r_2 = \frac{2b_1 - b_2}{7} \qquad r_1 = \frac{b_1 + 3b_2}{7}$$

Por lo cual dando valores arbitrarios a b_1 y b_2 se obtienen los valores para r_1 y r_2 que multiplicados por v_1 y v_2 pueden expresar como combinación lineal a b_1 y b_2.

Para la independencia lineal, se tiene que:

$$r_1 + 3r_2 = 0$$
$$r_2 = 0$$

Lo que implica $r_1 = r_2 = 0$ por lo tanto S es una base de R^2.

Teorema 5.7
Sea $S= \{v_1, v_2, \cdots, v_n\}$ una base para el espacio vectorial V, y sea $v \in V$. Los coeficientes en la representación.

$$v= r_1 v_1 + \cdots + r_n v_n$$

Son únicos.

Demostración
Suponga que se tienen las dos representaciones:

$$v= a_1 v_1 + \cdots + a_n v_n$$
$$v= b_1 v_1 + \cdots + b_n v_n$$

Para v; se demostrará que los coeficientes son iguales. Para ello, formamos $v+(-v)$ que es igual a cero y se combinan los términos para obtener:

$$0= (a_1-b_1)v_1 + \cdots + (a_n-b_n)v_n$$

Como S es una base, es un conjunto linealmente independiente, por lo tanto:

$$a_1-b_1=a_2-b_2= \cdots =a_n-b_n=0$$

Lo que implica:

$$a_1=b_1, \quad a_2=b_2, \cdots, \quad a_n=b_n$$

Ejemplo 5.24
Sean $v_1= (1, 2, 1)$, $v_2= (2, 9, 0)$ y $v_3= (3, 3, 4)$. Demuestre que el conjunto $S= \{v_1, v_2, v_3\}$ es una base de R^3.

Solución
Entonces se tiene que demostrar que para cualquier vector $b = (b_1, b_2, b_3)$ la ecuación:

$$r_1 v_1 + r_2 v_2 + \cdots + r_3 v_3= b$$

Siempre tiene solución, y que la única solución de:

$$r_1 v_1 + r_2 v_2 + \cdots + r_3 v_3 = 0$$

Es la trivial, esto es $r_1 = r_2 = r_3 = 0$.

De estos dos sistemas podemos concluir, que la matriz de coeficientes asociada

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix}$$

es invertible, esto es que existe solución única para $Ar=0$ y para $Ar=b$. Que A es invertible se puede verificar con su determinante, si det $A \neq 0$ entonces A es invertible.

$$det\ A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix} = -1$$

Por lo tanto, S es una base de R^3.

Teorema 5.8

Sean $v_1, v_2, \cdots, v_n$ vectores en R^n las siguientes condiciones son equivalentes:

1. Los vectores son linealmente independientes.
2. Los vectores generan todo R^n.
3. La matriz A que tiene estos vectores como vectores columna es invertible.

Ejemplo 5.25

Sea el conjunto $S = \{M_1, M_2, M_3\}$ tal que:

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Demuestre que S es una base del espacio vectorial M_{22} de las matrices de 2×2.

Solución

Para comprobar que S genera a M_{22}, observe que, dada una matriz arbitraria

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Se puede expresar como:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = aM_1 + bM_2 + cM_3 + dM_4$$

Es decir, una combinación lineal de M_1, M_2, M_3, y M_4.

Para verificar que S es linealmente independiente hacemos:

$$aM_1 + bM_2 + cM_3 + dM_4 = 0$$

O bien:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Entonces

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Por lo tanto, $a=b=c=d=0$ y S es linealmente independiente.

Ejemplo 5.26

Muestre que el conjunto $S = \{(1, 2), (3, -1), (1, 0)\}$ no es una base de R^2.

Solución

El conjunto S es linealmente dependiente porque:

$$(1, 2) + 2(3, -1) - 7(1, 0) = (0, 0)$$

Por lo cual S no es una base de R^2.

Teorema 5.9: reducción de un conjunto generador a una base
Si $S = \{v_1, \cdots, v_m\}$ es un conjunto de vectores no nulos que genera a un subespacio W de un espacio vectorial V, entonces algún subconjunto de S es una base para W. Dicha base se puede encontrar excluyendo de S aquellos vectores que son combinaciones lineales de predecesores.

Demostración
Si S es un conjunto linealmente independiente, entonces por definición S es una base para W.

Si S es linealmente dependiente, uno de los vectores puede escribirse como una combinación lineal de los otros. Suponga que v_m es dicho vector. Se afirma que $S' = \{v_1, \cdots, v_{m-1}\}$ aún genera a W. Para comprobar esto, sea x un elemento de W con:

$$x = c_1 v_1 + \cdots + c_{m-1} v_{m-1} + c_m v_m$$

Pero $v_m = d_1 v_1 + \cdots + d_{m-1} v_{m-1}$ por lo que esta expresión se puede sustituir en la combinación lineal anterior para obtener:

$$x = c_1 v_1 + \cdots + c_{m-1} v_{m-1} + c_m (d_1 v_1 + \cdots + d_{m-1} v_{m-1}) =$$
$$(c_1 + c_m d_1) v_1 + \cdots + (c_{m-1} + c_m d_{m-1}) v_{m-1}$$

Por lo tanto, S' genera a W. Si S' es linealmente independiente S' es una base para W. Si S' es linealmente dependiente uno de los vectores en S' es una combinación lineal de los demás y se procede a efectuar el mismo procedimiento. En esta forma debe llegarse al fin a un conjunto linealmente independiente que genere a W y que, por lo tanto, sea una base de W.

Se llega así al siguiente resultado fundamental:

Cualquier espacio vectorial finito, generado por un conjunto de vectores no nulos, tiene por lo menos una base.

Ejemplo 5.27
Considere el subespacio W de R^3 generado por el conjunto de vectores.

$S = \{(2, 1, 3), (-1, 2, 0), (1, 8, 6)\}$ encuentre una base de W

Solución
Comenzamos por examinar el segundo vector, como (−1, 2, 0) no es múltiplo de (2, 1, 3) entonces (−1, 2, 0) se queda en la lista. Procedemos con el vector (1, 8, 6) y verificamos si puede expresarse en la forma:

$$(1, 8, 6) = r_1(2, 1, 3) + r_2(-1, 2, 0)$$

Para escalares r_1, r_2, este sistema de ecuaciones se puede escribir en forma matricial como:

$$\left[\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 2 & 8 \\ 3 & 0 & 6 \end{array}\right]$$

Para encontrar la solución en r_1 y r_2 procedemos a efectuar eliminación gaussiana:

$$\left[\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 2 & 8 \\ 3 & 0 & 6 \end{array}\right] \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 2 & -1 & 1 \\ 3 & 0 & 6 \end{array}\right] \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 0 & -5 & -15 \\ 0 & -6 & -18 \end{array}\right] \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right]$$

Obtenemos la solución $r_2=3$ y $r_1=2$ y concluimos que (1, 8, 6) está en el conjunto:

$S' = \{(2, 1, 3), (-1, 2, 0)\}$ donde S' es una base de W

Tal como se ha presentado hasta ahora, la técnica de exclusión parece potencialmente laboriosa. Por ejemplo, para aplicarla a una lista $v_1, v_2, \cdots, v_6$ parece que debemos probar si tienen soluciones para cinco ecuaciones vectoriales.

$$v_2 = r_1 v_1, \qquad v_3 = r_1 v_1 + r_2 v_2$$

Y así sucesivamente.

En el espacio vectorial R^n, cada una de estas ecuaciones conducirá a un sistema lineal, según se ilustró en el ejemplo anterior. En realidad, se pueden resolver cinco sistemas lineales al mismo tiempo, reduciendo una sola matriz, según se enuncia en el recuadro.

Técnicas de exclusión en R^n

La eliminación de los vectores distintos de cero que son combinación lineal de los predecesores en una lista $v_1, v_2, \cdots, v_k$ se puede lograr como sigue:

1) Formar la matriz A con v_j como el j-ésimo vector columna.
2) Reducir la matriz A a la forma escalonada por filas como en el método de Gauss.
3) Los vectores de las columnas de A que dan lugar a pivotes (distintos de cero) deben permanecer y generarán todo el espacio columna de A. Esto es, se excluyen de la lista aquellos vectores de las columnas de A que no proporcionen pivotes distintos de cero.

Explicaremos la validez de esta técnica en el siguiente ejemplo.

Ejemplo 5.28

Sea W el subespacio de R^5generado por:

$v_1 = (1, -1, 0, 2, 1)$ $\quad$ $v_2 = (2, 1, -2, 0, 0)$

$v_3 = (0, -3, 2, 4, 2)$ $\quad$ $v_4 = (3, 3, -4, -2, -1)$

$v_5 = (2, 4, 1, 0, 1)$ $\quad$ $v_6 = (5, 7, -3, -2, 0)$

Use la técnica de exclusión para intentar acortar la lista v_i con $i = 1, \cdots, 6$ de generadores de W.

Solución

Se reduce la matriz que tiene v_j como j-ésimo vector columna y se tiene una forma escalonada.

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ -1 & 1 & -3 & 3 & 4 & 7 \\ 0 & -2 & 2 & -4 & 1 & -3 \\ 2 & 0 & 4 & -2 & 0 & -2 \\ 1 & 0 & 2 & -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ 0 & 3 & -3 & 6 & 6 & 12 \\ 0 & -2 & 2 & -4 & 1 & -3 \\ 0 & -4 & 4 & -8 & -4 & -12 \\ 0 & -2 & 2 & -4 & -1 & -5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ 0 & 1 & -1 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ 0 & 1 & -1 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

De la última matriz vemos que el subespacio W está generado por las columnas que contienen pivotes distintos de cero y corresponden a v_1, v_2y v_5 en la matriz original. Por lo tanto, se concluye que:

$$W = gen\,\{v_1, v_2, v_5\}$$

Definición 5.10

Un espacio vectorial V tiene dimensión finita si hay un subconjunto finito de vectores de V, $\{v_1, v_2, \cdots, v_n\}$ que formen una base de V (la dimensión es igual al número de vectores en una base de V). Si no existe tal subconjunto, V tiene *dimensión infinita* (además se convendrá en que el espacio trivial tiene dimensión finita, aun cuando no tenga conjuntos linealmente independientes y por consecuencia, no tenga ninguna base).

La dimensión de un espacio vectorial V se denota por ***dim V.***

Teorema 5.10

1. Si $S= \{v_1, \cdots, v_m\}$ es un conjunto de n vectores linealmente independientes en un espacio V de n dimensiones, entonces S es una base de V.
2. Si $S= \{v_1, \cdots, v_m\}$ es un conjunto de n vectores que generan un espacio V de n dimensiones, entonces S es una base de V.
3. Si $S= \{v_1, \cdots, v_m\}$ es un conjunto linealmente independiente en un espacio V de n dimensiones y $r < n$, entonces a S se le pueden agregar vectores hasta formar una base de V, es decir, existen vectores $v_{r+1}, \cdots, v_n$ tales que: $\{v_1, v_2, \cdots, v_r, v_{r+1}, \cdots, v_n\}$ es una base de V

Este teorema dice que, si se conoce la dimensión n de un espacio vectorial V de dimensión finita, entonces los incisos 1 y 2 simplifican el problema de determinar si un conjunto S de n vectores es una base. Para mostrar que S es una base, basta con mostrar que S genera a V, o que S es linealmente independiente. Es claro que si no se conoce la dimensión de V se deben probar ambas cosas. La utilidad del inciso 3 se verá más adelante.

Teorema 5.11

Si $S = \{v_1, \cdots, v_n\}$ es una base de V, entonces:

a) Cualquier conjunto de m vectores con $m > n$ es linealmente dependiente, y por lo tanto no es una base para V.

b) Cualquier conjunto de m vectores con $m < n$ no puede generar a V, por lo tanto, no es una base de V.

Este teorema simplemente dice que el número de vectores es único en una base.

Demostración

a) Sea $T = \{w_1, w_2, \cdots, w_n, w_{n+1}\}$. Es decir, T contiene exactamente $n+1$ vectores. Se demostrará que T no puede ser una base probando que T es linealmente dependiente. Para ello considere:

$$c_1 w_1 + c_2 w_2 + \cdots + c_n w_n + c_{n+1} w_{n+1} = 0 \tag{5.6}$$

Pero cada w_k puede escribirse como:

$$w_k = a_{1k} v_1 + a_{2k} v_2 + \cdots + a_{nk} v_n \qquad k = 1, 2 \cdots, n+1$$

Ya que S genera a V. Sustituyendo esto en la ecuación (5.6) se tiene:

$$\sum_{k=1}^{n+1} c_k w_k = \sum_{k=1}^{n+1} c_k \sum_{j=1}^{n} a_{jk} v_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{n+1} a_{jk} c_k \right) v_j = 0$$

Desarrollando se tiene:

$$\begin{aligned}
a_{11} c_1 v_1 + a_{12} c_2 v_2 + \cdots + a_{1,n+1} c_{n+1} v_n &= 0 \\
a_{21} c_1 v_1 + a_{22} c_2 v_2 + \cdots + a_{2,n+1} c_{n+1} v_n &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1} c_1 v_1 + a_{n2} c_2 v_2 + \cdots + a_{n,n+1} c_{n+1} v_n &= 0
\end{aligned}$$

Observe que este sistema homogéneo tiene menos ecuaciones que incógnitas, por lo que existe entonces solución no trivial para $c_1, c_2, \cdots, c_{n+1}$. Esto quiere decir que T es linealmente dependiente. Ahora suponga que T contiene más de $n+1$ vectores. Sea T' un subconjunto de $n+1$ vectores, tal como se acaba de demostrar, debe ser linealmente dependiente. Como $T'T$, el conjunto T contiene un subconjunto linealmente dependiente y es entonces un conjunto linealmente dependiente.

a) Suponga que T contiene $n-1$ vectores y que genera a V. Entonces por el teorema 5.8, T debe contener una base S^* para V. Si S^* contiene r vectores, entonces deberá ser $r \leq n-1$. Como S^*es una base y S tiene $r+1$ o más vectores, se infiere de la parte a) que S es linealmente dependiente. Lo que contradice que S sea una base.

Teorema 5.12

Cualesquiera dos bases de un espacio vectorial de dimensión finita tienen el mismo número de vectores.

Demostración

Sean $S= \{v_1, \cdots, v_n\}$ y $S'= \{w_1, w_2, \cdots, w_m\}$ dos bases de un espacio vectorial V de dimensión finita. Dado que S es una base y S' es un conjunto linealmente independiente por el teorema 5.9 se afirma que $m \leq n$. De igual manera puesto que S' es una base y S es un conjunto linealmente independiente se tiene $n \leq m$ lo que implica que $m=n$.

Ejemplo 5.29 (Base canónica)

Muestre que R^n tiene la base canónica

$E = \{e_1, e_2, \cdots, e_n\}$ donde $e_j = (0, 0, \cdots, 1_{j\text{-}componente}, 0, \cdots, 0)$

Solución

Para el espacio generado considere cualquier $x = (x_1, x_2, \cdots, x_n) \in R^n$ observe que:

$$x = x_1 (1, 0, \cdots, 0) + x_2 (0, 1, \cdots, 0) + \cdots + x_n (0, \cdots, 0, 1) =$$
$$x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

Para el caso de la independencia lineal considere:

$$c_1 e_1 + c_2 e_2 + \cdots + c_n e_n = (0, 0, \cdots, 0)$$

De donde se ve que $c_1 = c_2 = \cdots = 0$. Por lo tanto, E es una base de R^n y $dim\ R^n=n$

Esto es razonable ya que se tiene:

R – recta	objeto unidimensional
R^2– plano	objeto bidimensional
R^3– espacio	objeto tridimensional

Ejemplo 5.29

Demuestre que M_{23} tiene dimensión 6.

Solución

Una base que es la canónica de M_{23} es la siguiente:

$$S=\left\{\begin{matrix}\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&0&1\\0&0&0\end{bmatrix}\\ \begin{bmatrix}0&0&0\\1&0&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&0&1\end{bmatrix}\end{matrix}\right\}$$

Por lo que $dim\ M_{23}$ = número de vectores en $S = 6$.

Ahora que ya sabemos qué es una base de un espacio vectorial, podemos plantear el segundo problema fundamental del álgebra lineal.

EL PROBLEMA DE LA BASE

Sea V un espacio vectorial. El problema de la base puede presentarse en una de las siguientes formas:

Problema 1. Construir una base para V. Tomando los vectores de V.
Problema 2. Dado un conjunto S de vectores en V, construir una base para V añadiendo o bien eliminando algunos (pero no todos) los vectores de S, o ambas cosas.

Se puede resolver esto planteando lo siguiente:

Problema 1. Si puede escogerse un conjunto de vectores de V que genere V entonces eliminando vectores dependientes se obtendrá una base para V.
Problema 2. Si el conjunto dado S genera V, se procede como en el problema 1. Si no, se aumenta S añadiendo más vectores hasta que se obtenga un conjunto generador y linealmente independiente.

Ejemplo 5.30 (Segundo problema de la base)
Sea $S = \{(1, 0, 3), (2, 1, 4)$, encuentre una base T para R^3 que contenga a S.

Como R^3 tiene dimensión 3, entonces T debe contener 3 vectores. El conjunto S, como está, ya es linealmente independiente por lo que solo hace falta añadir un vector a S, de tal forma que este nuevo vector sea linealmente independiente con los dos ya existentes.

Solución 1: (prueba y error)
Como $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ es una base para R^3 entonces verificamos la independencia lineal de cada uno de estos vectores con los de S. Es decir, si existen r_1, r_2y r_3 tales que:

$$r_1 (1, 0, 3) + r_2 (2, 1, 4) + r_3 (1, 0, 0) = (0, 0, 0)$$

Lo que implica que:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 4 & 0 & 0 \end{array}\right] \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Así los vectores son linealmente independientes y forman una base:

$$T = \{(1, 0, 3), (2, 1, 4), (1, 0, 0)\}$$

Solución 2:
Otra forma de resolverlo es la siguiente:

$$\begin{aligned} gen\,\{(1, 0, 3), (2, 1, 4)\} &= \{\, x \mid x = a\,(1, 0, 3) + b\,(2, 1, 4)\} \\ &= \{\, x \mid x = (a + 2b,\ b,\ 3a + 4b)\} \end{aligned}$$

Por lo tanto, hay que procurar que el nuevo vector no sea de la forma:

$$(a + 2b, b, 3a + 4b).$$

Para ello se supone que el nuevo vector es (x_1, x_2, x_3) y se hace que la ecuación

$$(a + 2b, b, 3a + 4b) = (x_1, x_2, x_3)$$

no tenga solución para *a* y *b*. Igualando las componentes se obtienen las ecuaciones cuya representación matricial es:

$$\left[\begin{array}{cc|c} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 3 & 4 & x_3 \end{array}\right]$$

Lo que se reduce a:

$$\left[\begin{array}{cc|ccc} 1 & 2 & x_1 & & \\ 0 & 1 & x_2 & & \\ 3 & 4 & x_3 & -3x_1 & +2x_2 \end{array}\right]$$

Por lo tanto, si x_1, x_2 y x_3 se eligen de tal manera que $x_3 - 3x_1 + 2x_2^{1}$ 0 se tendrá un vector que no está en *gen* {1, 0, 3), (2, 1, 4)}, en tal caso $x = (0, 1, 0)$ es una elección aceptable y:

$$T = \{(1, 0, 3), (2, 1, 4), (0, 1, 0)\}$$

Es una base de R^3.

Solución 3

Si el tercer vector (x_1, x_2, x_3) hiciera que $\{(1, 0, 3), (2, 1, 4), (x_1, x_2, x_3)\}$ fuera linealmente dependiente, entonces:

$$\det \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ x_1 & x_2 & x_3 \end{bmatrix} \neq 0$$

Calculando el determinante se tiene:

$x_3 + 2x_2 - 3x_1 + 0$

Que es la misma condición obtenida mediante la solución 2. Algunos terceros vectores posibles son (1, 0, 0), (0, 1, 0), (0, 1, 1) o (1, 1, 0) en realidad hay un número infinito de posibilidades.

5.7. RESUMEN

1. Un espacio vectorial es un conjunto no vacío V de objetos llamados vectores junto con reglas para sumar dos vectores cualesquiera v y w de V y para multiplicar cualquier vector v de V y cualquier escalar r de R. Se requiere que V sea cerrado bajo esta suma de vectores y multiplicación por un escalar, de modo que $v+w$ y rv estén en V. Es más, los axiomas siguientes se deben satisfacer para todos los vectores u, v y w de V y todos los escalares r y s de R.

2. Para todos los vectores v y w de R^n, tenemos:

   Desigualdad de Schwarz: $|v \cdot w| \leq \|v\| \cdot \|w\|$

   Desigualdad triangular: $|v+w| \leq \|v\| + \|w\|$

3. Un subconjunto S de un espacio vectorial V es un subespacio de V si y solo si, es no vacío y satisface las dos propiedades de cierre:

   $v + w \in S$ para todos los vectores v y $w \in S$
   $rv \in S$ para todos los vectores $v \in S$ y escalares $r \in R$

4. El espacio generado por $\{v_1, v_2, \cdots, v_n\}$ es el conjunto de todas las combinaciones posibles de $v_1, v_2, \cdots, v_n$ y se denota:

   $gen\ (v_1, v_2, \cdots, v_n) = \{r_1\, v_1 + r_2\, v_2 + \cdots + r_n v_n\ r_i \in R\ i=1, 2, \cdots, n\}$

5. Sea V un espacio vectorial, un conjunto de vectores $\{v_1, v_2, \cdots, v_k\}$ de V es linealmente dependiente si existe una relación de dependencia:

   $r_1 v_1 + r_2 v_2 + \cdots + r_n v_n = 0$, para algún $r_i \neq 0$

   El conjunto es linealmente independiente si no existe dicha relación de dependencia, de modo que una combinación lineal de los v_i es el vector cero solo si todos los coeficientes son cero.

6. Una lista finita de vectores distintos de cero en V forma un conjunto linealmente independiente si, y solo si, ningún vector de la lista se puede expresar como combinación lineal de sus predecesores.

7. Un conjunto de vectores $\{v_1, v_2, \cdots, v_k\}$ es una base para un subespacio S de un espacio vectorial V si es linealmente independiente y constituye un conjunto generador para S.

8. Cualquier subconjunto generador finito de un subespacio S, distinto de cero, de un espacio vectorial V se puede reducir, de ser necesario, a una base para S quitando los vectores cero y excluyendo después aquellos que son combinaciones lineales de predecesores.

9. Las siguientes afirmaciones son equivalentes para n vectores de R^n.

   a) Los vectores forman una base para R^n
   b) Los vectores son linealmente independientes
   c) Los vectores generan R^n
   d) Una matriz que tenga a los vectores como columnas es invertible

10. Sea S un subespacio de dimensión finita de un espacio vectorial V. Todas las bases de S tienen el mismo número de elementos. Este número, $dim(S)$, se llama dimensión de S.

5. 8. APLICACIONES DE ESPACIOS VECTORIALES

La aplicación que se presenta en esta sección tiene como objetivo presentar otro paradigma de los usos de este tema, con el objetivo de ampliar la perspectiva del lector sobre el álgebra lineal.

LOS ESPACIOS VECTORIALES, EL AMARILLO, EL ROJO Y EL AZUL

(Tomado de Suma 37, junio 2011, pp.75-82, Miguel Ángel Moreno Redondo)

Como sabemos, el amplio espectro de colores que nuestra retina puede llegar a recibir, varía desde el amarillo, el azul, el rosa, el violeta y decenas de colores y tonalidades distintas, que por raro que parezca, se pueden obtener a partir de los pigmentos de los tres colores primarios: amarillo, azul y rojo. A título de ejemplo sirva que el color verde se obtiene mezclando el amarillo y el azul, el naranja mezclando el rojo y el amarillo, etc. Por otra parte, si solo utilizamos los colores blanco y negro, podremos obtener todas las tonalidades posibles de grises, además de las dos siguientes obviedades: el blanco lo podemos obtener a partir del blanco, y el negro a partir del negro.

Entendido el párrafo anterior no resulta difícil comprender, que al genial pintor sevillano Velázquez le habrían bastado los tres colores primarios para obtener todos los colores del cuadro *Las meninas* y que sin uno solo de ellos le habría sido totalmente imposible. Por otra parte, y como un ejercicio propuesto a la intuición, plantearemos la siguiente pregunta: ¿qué dos colores le habrían bastado a Picasso para pintar el *Guernica*?

La interpretación que haremos será la de establecer una relación biunívoca entre los elementos: vector y color, y la que resulta aún más interesante: la existente entre combinación lineal y mezcla de colores. Una vez establecidas estas identificaciones, volveremos a escribir en términos de colores los distintos enunciados que acostumbramos a utilizar. Por último, y sin lugar a dudas, principal atracción del artículo, veremos una lista de trece ejemplos que plasman la utilidad de este nuevo enfoque.

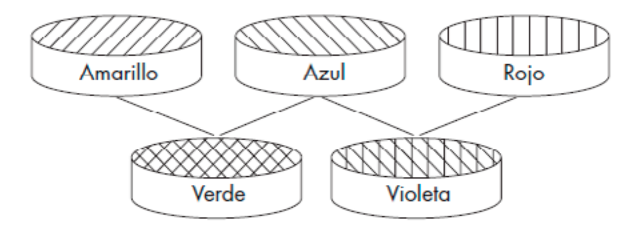


Figura 1. Los colores primarios amarillo, azul y rojo, nos proporcionan entre otros el verde y el violeta, tras la mezcla de sus pigmentos

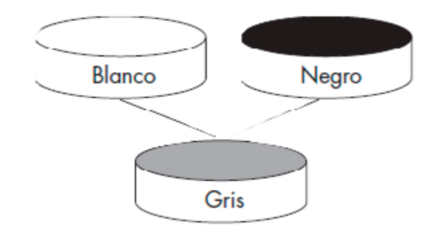


Figura 2. Los colores blanco y negro nos proporcionan una amplia gama de grises, entre ellos el gris claro (calificativo que alcanza al mezclarse en mayor proporción el blanco respecto al negro)

Analogías

A continuación, presentamos el siguiente cuadro de equivalencias:

Terminología formal	*Terminología informal*
Espacio vectorial	Cuadro
E	Las meninas
V	El Guernica
Vector	Color
Vectores	Colores
Combinación lineal	Mezcla

Como podemos observar todo término de la columna de la izquierda tiene su respectivo término en la columna de la derecha, al que podríamos llamar además de término informal, término dual. Cada término dual gozará de total significado en la construcción de los enunciados que más tarde veremos. Sirvan como aperitivo las siguientes observaciones:

Al igual que un vector forma parte de la estructura de espacio vectorial, un color formará parte del espectro de colores que baña cuanto nos rodea.

» No hay que sorprenderse tampoco de la analogía entre E y el cuadro de Velázquez *Las meninas* o entre V y el *Guernica*, pues al igual que la combinación lineal de las sucesivas potencias de x nos proporciona el conocido

espacio vectorial de los polinomios, la mezcla de colores nos ofrece la maravillosa posibilidad de pintar *Las meninas*.

» Por último, sería conveniente recordar que mientras *Las meninas* fue pintado con una amplia gama de colores, Picasso utilizó en el *Guernica* el blanco, el negro y distintas tonalidades de grises.



Figura 3. Las meninas, cuadro pintado por D. Velázquez en 1656. Los colores que lo componen son el amarillo, el azul y el rojo, además de una amplia gama de colores

Dualidades

Cada enunciado de los que a continuación se presentan expresados de una manera formal en términos de espacios vectoriales, aparece con su respectivo enunciado expresado en términos de colores. Si bien el primero, a cualquier iniciado en el tema le resultará algo confuso, el segundo lo percibirá de una forma nítida y clara.

Teorema 1. Sistema linealmente dependiente

» Dado un sistema de vectores S, decimos que los vectores de S son *linealmente dependientes* si, y solo si, existe un vector de S, que es combinación lineal de los demás.

» Dado un sistema de colores S, decimos que los colores de S son *linealmente dependientes* si, y solo si, existe un color de S que es mezcla de los demás.

Teorema 2. Sistema linealmente independiente

» Sea S un sistema de vectores. Los vectores de S son *linealmente independientes* si, y solo si, ningún vector de S es combinación lineal de los demás.

» Sea S un sistema de colores. Los colores de S son *linealmente independientes* si, y solo si, ningún color de S es mezcla de los demás.

Definición. Sistema generador

» Decimos que S es un *sistema* de vectores *generador* del espacio vectorial E si todo vector de E es combinación lineal de los vectores de S.

» Decimos que S es un *sistema* de colores *generador* del cuadro *Las meninas* si todo color de *Las meninas* es mezcla de los colores de S.

Definición. Base

» Decimos que S es una base de vectores del espacio vectorial E si los vectores de S son linealmente independientes *y generadores del espacio vectorial E.*

» Decimos que S es una base de colores del cuadro Las meninas si los colores de S son linealmente independientes y generadores del cuadro Las meninas.

Definición. Dimensión

» Decimos que un espacio vectorial tiene dimensión n si n es el número de vectores de una cualquiera de sus bases.

» Decimos que un cuadro tiene dimensión n si n es el número de colores de una cualquiera de sus bases

» No nos costaría imaginar, lo que podría llegar a ser la interpretación de combinación lineal. La expresión $v= \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n$ con $\lambda_i \geq 0$ puede significar que los colores $e_1, e_2, \cdots, e_n$ deben mezclarse en cantidades $\boldsymbol{\lambda}_1, \lambda_2, \cdots, \lambda_n$ respectivamente para obtener el color v. En particular, veamos cómo podemos obtener algunos colores:

- El verde se podría obtener mezclando 1 parte de amarillo y 1 parte de azul.
- El verde muy oscuro se podría obtener mezclando 1 parte de amarillo y 3 partes de azul.
- El gris claro se podría obtener mezclando 3 partes de blanco y 1 parte de negro.

Ejemplo 1

Los colores blanco y negro son un sistema generador del cuadro *Guernica*, ya que con esos dos colores se pueden obtener todas las tonalidades posibles de grises, blanco y negro para poder pintarlo.

Ejemplo 2

Los colores blanco y negro son independientes, ya que el blanco no se puede obtener a partir del negro y viceversa (teorema 2).

Ejemplo 3

Los colores blanco y negro son base de los colores que se utilizan para pintar el *Guernica*, pues cualquier color del cuadro se puede obtener como mezcla del blanco y el negro (colores generadores) y además, ninguno de ellos dos se puede obtener a partir del otro (Teorema 2, independencia). Por tanto: ni faltan, ni sobran.



Figura 4. Guernica, cuadro pintado por P. Picasso en 1937.
Los colores que lo componen son el blanco, el negro y distintas tonalidades de grises

Ejemplo 4

El amarillo y el azul son colores linealmente independientes, pues ni el amarillo se puede obtener a partir del azul, ni el azul a partir del amarillo.

Ejemplo 5

El amarillo, el azul y el verde no son colores linealmente independientes, pues el verde se obtiene mezclando el amarillo y el azul (el verde es combinación lineal del amarillo y el azul).

Ejemplo 6

El amarillo, el azul y el rojo son linealmente independientes pues ninguno de ellos se puede obtener a partir de los demás.

Ejemplo 7

El amarillo, el azul, el rojo y el verde es un sistema linealmente dependiente pues, en particular, el verde se puede obtener a partir del verde o bien del amarillo y el azul (como vemos, no hay unicidad de coordenadas).

Ejemplo 8

El amarillo, el azul y el rojo (colores primarios), forman un sistema generador de los colores del cuadro *Las meninas*, pues cualquier color se puede obtener mezclando debidamente estos tres (combinación lineal).

Ejemplo 9

El amarillo y el azul no forman un sistema generador de los colores del cuadro *Las meninas*, pues en particular el rojo, no se puede obtener mezclando estos dos.

Ejemplo 10

El amarillo, el azul, el rojo y el verde forman un sistema generador de los colores del cuadro *Las meninas*, pues cualquier color se puede obtener mezclando debidamente estos cuatro y en particular los tres primeros (combinación lineal).

Ejemplo 11

El amarillo, el azul y el rojo forman una base de los colores del cuadro *Las meninas,* ya que son linealmente independientes y además generadores (ejemplos 6 y 8).

Ejemplo 12

El cuadro *Las meninas* tiene dimensión 3, ya que 3 es el número de colores de la base formada por los colores amarillo, azul y rojo (ejemplo 11).

Ejemplo 13

El cuadro *Guernica* tiene dimensión 2, ya que 2 son los colores de la base formada por los colores blanco y negro (ejemplo 3).

5.9. NOTAS HISTÓRICAS

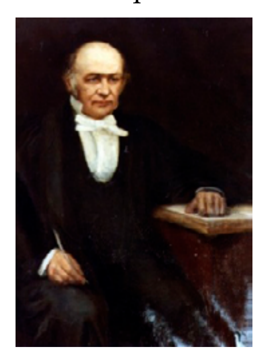


La idea de un espacio n–dimensional para $n>3$ se aceptó gradualmente durante el siglo XIX; por tanto, es difícil señalar una primera "invención" de este concepto. Entre los diversos usos tempranos de este concepto está su presencia en un trabajo del matemático ruso **Mikhail Ostrogradskii** (1801-1862) sobre el teorema de la divergencia, escrito en 1836.

En los tratados geométricos de **Hermann Grassmann** (1809-1877) al principio de la década de 1840, y en un breve artículo de Arthur Cayley, en 1846. Lamentablemente, los dos primeros autores fueron virtualmente ignorados en vida. En particular, el trabajo de Grassman era bastante filosófico y muy difícil de leer.

Arthur Cayley (Richmond, Reino Unido, 16 de agosto de 1821-Cambridge, 26 de enero de 1895) fue un matemático británico. Es uno de los fundadores de la escuela británica moderna de matemáticas puras. El artículo de Cayley simplemente decía que se podían generalizar ciertos resultados a dimensiones mayores que tres "sin recurrir a ninguna noción metafísica respecto a la posibilidad de un espacio de cuatro dimensiones".

Sir William Rowan Hamilton (1805-1865), en una carta de 1841, también observó que "debe de ser posible, de una u otra manera, introducir no solo



ternas sino poliadas (polyplets), de modo que en algún sentido se satisfaga la ecuación simbólica $a= (a_1,a_2, \cdots, a_n)$ donde a es un símbolo indicativo de un pensamiento (complejo) y a_1, a_2, $\cdots$, a_n denotan n números reales, positivos o negativos".



Aunque los vectores de R^n fueron tratados por los matemáticos y físicos durante la segunda mitad del siglo XIX (como sucedió con otros objetos que hoy consideramos vectores), hasta la aparición del tratado de **Hermann Weyl** (1885-1955) *Space-Time Matter* en 1918, no se imprimió una definición axiomática y abstracta de un espacio vectorial. Weyl escribió este libro como introducción

a la teoría general de la relatividad de Einstein; en el capítulo 1 analizó la naturaleza del espacio euclidiano y como parte de ese análisis, formuló lo que ahora son los axiomas comunes de un espacio vectorial. Como trató solo los espacios V de dimensión finita, incluyó el axioma de dimensionalidad: para algún número entero h hay h vectores linealmente independientes en V, pero todo conjunto de $h+1$ vectores es linealmente dependiente.

Los axiomas de espacio vectorial se conocían desde hacía años, pero en general se demostraban como consecuencia de otras definiciones de vectores. Por ejemplo, aparecen en el breve texto en italiano de **Giuseppe Peano** (1858-1932) *Cálculo geométrico* (1888), donde se explica el trabajo de Hermann Grassmann.



La desigualdad de Schwarz se debe independien-temente al trabajo de Augustin-Louis Cauchy (1789-1857) y el de **Hermann Amandus Schwarz** (1843-1921) y Viktor Yakovlevich Bunyakovsky (1804-1889). Primero se formuló como un teorema sobre coordenadas en un apéndice al texto de Cauchy de 1821 para su curso de análisis en la *Ecole Polytechnique* como sigue:

$$|aa+a'\alpha'+a''\alpha''+\cdots| \le \sqrt{a^2+a'^2+a''^2+\cdots}\;\sqrt{\alpha^2+\alpha'^2+\alpha''^2+\cdots}$$

Bunyakovsky demostró en 1859 la desigualdad para funciones, esto es, formuló el resultado

$$\left[\int_a^b f(x)\,g(x)dx\right]^2 \le \int_a^b f^2(x)dx \int_a^b g^2(x)dx$$

Donde la integral del producto de las funciones f y g se puede considerar como el producto interior de las funciones continuas en $[a, b]$. Bunyakovsky fue vicepresidente de la Academia de Ciencias de San Petersburgo desde 1864 hasta su muerte. En 1875 la Academia estableció un premio matemático en su nombre, como reconocimiento a sus 50 años de enseñanza e investigación.

Schwarz formuló la desigualdad en 1884. En su caso, los vectores eran funciones Φ, Ω de dos variables en una región T del plano, y el producto interior de estas funciones estaba dado por la doble integral de su producto, donde se supone que esta integral existe. Entonces la desigualdad establece que:

$$\left|\iint_T \Phi\Omega dxdy\right| \leq \sqrt{\iint_T \Phi^2\, dxdy} * \sqrt{\iint_T \Omega^2\, dxdy}$$

Schwarz era el principal matemático de Berlín alrededor del cambio de siglo, el trabajo donde aparece la desigualdad está dedicado a una cuestión sobre superficies minimales.

5.10. EJERCICIOS PROPUESTOS

1. ¿Cuáles de los siguientes subconjuntos de R^3 son subespacios?
   a) El plano de los vectores con primera componente b_1 igual a cero.
   b) El plano de los vectores b con $b_1 = 1$.
   c) Los vectores b con $b_1b_2 = 0$ (esta es la unión de dos subespacios, el plano $b_1 = 0$ y el plano $b_2= 0$).
   d) El vector solitario $b =(0, 0, 0)$.
   e) Todas las combinaciones de los vectores $u =(1,1,0)$ y $v =(2,0,1)$.
   f) Los vectores (b_1,b_2,b_3) que satisfacen $b_3 - b_2 + 3b_1 = 0$.

2. Usar la desigualdad del triángulo para probar que:

   $\|v\text{-}w\| \leq \|v\| + \|w\|$

3. Pruebe si son falsos o verdaderos los siguientes postulados:
   a) Si $x_1, x_2,\cdots,x_n$ son vectores linealmente independientes, lo mismo es cierto de cualquier subconjunto de estos vectores.
   b) Si $x_1, x_2,\cdots,x_n$ son vectores linealmente dependientes, lo mismo es cierto de cualquier subconjunto de estos vectores.
   c) Si $x_1, x_2,\cdots,x_n$ generan el espacio vectorial x, entonces la dimensión de x es n.

Justifique su respuesta: pruebe si es cierto o proporcione contraejemplo si es falso.

4. ¿Es cierto que si v_1, v_2, y v_3 son linealmente independientes, entonces los vectores $w_1 = v_1 + v_2$, $w_2 = v_1 + v_3$ y $w_3 = v_2 + v_3$ son linealmente independientes?
5. En la definición de un Espacio Vectorial se exige que la suma y la multiplicación por un escalar satisfagan ciertas reglas o axiomas (ver notas de clase). Supongamos que a cada componente se le suma un "1" adicional, de modo que (3,1) + (5,0) es igual a (9,2) en lugar de (8,1). ¿Cuáles de las diez reglas se violan, si mantenemos sin cambio la multiplicación por un escalar?

6. Sea P el plano en el espacio R_3 dado por la ecuación $x + 2y + z = 6$. ¿Cuál es la ecuación del plano P_0 que pasa por el origen y es paralelo a P? ¿Son P y P_0 subespacios de R^3?

7. Decida acerca de la dependencia o independencia de los vectores.
   a) (1,1,2), (1,2,1) y (3,1,1).
   b) e_1-e_2, e_2-e_3, e_3-e_4 y e_4-e_1, donde e_1,e_2,e_3,e_4 son vectores de la base canónica de R^4.
   c) (1,1,0), (1,0,0), (0,1,1) y (x, y, z) para números x, y, z cualesquiera.

8. Para decidir si b está en el subespacio generado por el conjunto de vectores $w_1,\cdots,w_l$ considere a los vectores w_i (i=1,⋯, l) como las columnas de la matriz A. Intente resolver el sistema $Ax=b$. ¿Cuál es el resultado si se tiene:
   a) $w_1 =(1,1,0)$, $w_2 = (2,2,1)$, $w_3 =(0,0,2)$ y $b =(3,4,5)$.
   b) $w_1 =(1,2,0)$, $w_2 =(2,5,0)$, $w_3 =(0,0,2)$, $w_4 = (0,0,2)$ y con cualquier vector b.

9. Determine en cada caso si el conjunto de vectores dado es o no una base para el espacio indicado:
   a) $S = \{(-1, 1), (1, 2)\}$ para R^2.
   b) $T = \{(-1, 3, 4), (1, 5, -1), (1, 13, 2)\}$ para R^3.

10. En los siguientes incisos se dan un conjunto S y un espacio vectorial V. Encuentre una base de V que contenga a S.
    a) $V = R^3$ $S= \{(1, -1, 1), (0, 1, -1)\}$.
    b) $V = M_{22}$.

11. En los siguientes incisos se dan un conjunto S y un espacio vectorial V. Encuentre una base de V eliminando vectores de S.

a) $V = R^2 \quad \mathrm{S} = \{(1, -1), (1, 2), (3, 4)\}$.

b) $V = \mathrm{R}^3 \; S = \{(1, -1, 2), (4, -3, 7), (2, 0, 5), (1, 2, 6)\}$.

Capítulo 6
LOS CUATRO ESPACIOS FUNDAMENTALES

6.1. ESPACIOS ASOCIADOS A UNA MATRIZ

Cuando tenemos un sistema de n ecuaciones con n incógnitas, podemos resolverlo escribiéndolo en forma matricial y llevando su matriz de coeficientes a una matriz triangular mediante una eliminación gaussiana para finalmente proceder a resolver el sistema por sustitución regresiva.

Para el caso de sistemas con m ecuaciones y n incógnitas (o viceversa), se procede a efectuar una eliminación gaussiana, pero la matriz resultante, no es una matriz triangular superior estrictamente hablando, sino una matriz escalonada.

La forma final U es **"triangular superior"**, donde los pivotes no están en la diagonal principal y las entradas distintas de cero quedan en forma escalonada.

$$U=\begin{bmatrix} a & g & h & i \\ 0 & 0 & k & b \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Antes de estudiar todas las posibilidades de encontrar soluciones a este tipo de sistemas, conviene que estudiemos ciertos espacios asociados a las matrices; los resultados nos proporcionarán un procedimiento sencillo para elaborar bases reduciendo una matriz determinada a la forma escalonada, en consecuencia, tendremos la posibilidad de analizar soluciones a tales sistemas.

Definición 6.1

Considere la matriz:

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \vdots & \vdots & \ldots\ldots & \vdots \\ \vdots & \vdots & \ldots\ldots & \vdots \\ a_{m1} & a_{m2} & \ldots\ldots & a \end{bmatrix}$$

Y los vectores:

$$r_1 = (a_{11} \;\; a_{12} \cdots \;\; a_{1n})$$
$$r_2 = (a_{21} \;\; a_{22} \cdots \;\; a_{2n})$$
....
$$r_m = (a_{m1} \;\; a_{m2} \cdots \;\; a_{mn})$$

que se forman con los renglones de A, se denominan los **vectores renglón de *A***. El subespacio de R^n generado por los vectores renglón se llama **espacio de los renglones de *A***.

Ejemplo 6.1

Los vectores renglón de la matriz:

$$A=\begin{bmatrix} 2 & 1 & 0 \\ 3 & -1 & 4 \end{bmatrix}$$

Son $r_1 = (2, 1, 0)$ y $r_2 = (3, -1, 4)$

El siguiente teorema es de gran utilidad para los cálculos que se efectuarán posteriormente.

Teorema 6.1

Las operaciones elementales en los renglones no alteran el espacio de los renglones de una matriz.

Supongamos que se quiere elaborar una base para el subespacio de R^n generado por un determinado conjunto de vectores $\{v_1, v_2, \cdots, v_m\}$; entonces, si se forma la matriz B cuyos vectores renglón son $v_1, v_2, \cdots, v_m$, el problema se reduce a encontrar una base para el espacio de los renglones de B; por ejemplo, se puede obtener una base para el espacio generado por los vectores:

$$v_1 = (1, -2, 0, 0, 3)$$
$$v_2 = (2, -5, -3, -2, 6)$$
$$v_3 = (0, 5, 15, 10, 0)$$
$$v_4 = (2, 6, 18, 8, 6)$$

Encontrando una base para el espacio de los renglones de la matriz, entonces:

$$B=\begin{bmatrix}1 & -2 & 0 & 0 & 3\\ 2 & -5 & -3 & -2 & 6\\ 0 & 5 & 15 & 10 & 0\\ 2 & 6 & 18 & 8 & 6\end{bmatrix}$$

Por el teorema 6.1, se deduce que el espacio de los renglones de una matriz no se altera si la matriz se lleva a la forma escalonada.

$$U=\begin{bmatrix}1 & -2 & 0 & 0 & 3\\ 0 & 1 & 3 & 2 & 0\\ 0 & 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

U es la forma escalonada de B. Dado que el espacio de los renglones de B coincide con los de U, una base para el espacio de los renglones de U será también una base para el espacio de los renglones de B. Es fácil verificar que los vectores renglón de U, diferentes de cero, son linealmente independientes, por lo tanto, estos vectores forman una base para el espacio de los renglones de U y en consecuencia para el espacio de los renglones de B lo que nos conduce al teorema 6.2.

$r_1 = (1, -2, 0, 0, 3)$
$r_2 = (0, 1, 3, 2, 0)$
$r_3 = (0, 0, 1, 1, 0)$

Teorema 6.2
Los vectores renglón diferentes de cero, en la forma escalonada de una matriz A, constituyen una base para el espacio de los renglones de A.

Nota 1: Las columnas que contienen los pivotes son linealmente independientes.
Nota 2: La dimensión del espacio de los renglones es igual al rango de la matriz.

> ***Definición 6.2***
> El **espacio columna** de una matriz de orden $m \times n$ es el espacio generado por las columnas. Es un subespacio del espacio m-dimensional R^m. Si $m=n$ entonces tanto el espacio renglón como el espacio columna son subespacios de R^n, incluso pueden ser el mismo subespacio. Al espacio columna a menudo se le llama el recorrido de A: $R(A)$; similar a la idea de recorrido de una función, como se muestra en la figura 6.1

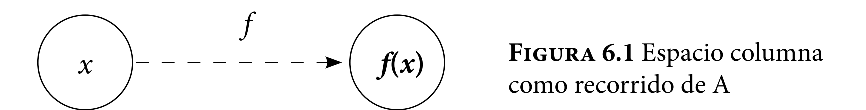


Figura 6.1 Espacio columna como recorrido de A

En este caso la función es $f(x)=Ax$, cuyo dominio consta de todas las x en R^n y el recorrido todos los posibles vectores Ax. Esto es, todas las b para las cuales puede resolverse $Ax=b$ es igual a todas las posibles combinaciones de las columnas (el recorrido es el espacio columna).

Ejemplo 6.2

Sea la matriz:

$$A=\begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

Las columnas de esta matriz son linealmente dependientes, ya que la segunda columna es tres veces la primera. Los renglones son también linealmente dependientes, el tercero es dos veces el segundo menos cinco veces el primero.

Si escalonamos la matriz A, los renglones distintos de cero deben ser linealmente independientes.

$$U=\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Además, si elegimos las columnas que contienen los pivotes, también son linealmente independientes (en este caso las columnas 1 y 3 son linealmente independientes).

Las cuatro columnas generan el espacio columna por definición, sin embargo, no son linealmente independientes; por lo tanto, proponemos que las columnas que contengan los pivotes distintos de cero sean una base del espacio columna.

Los dos primeros renglones (por el teorema 6.2) son una base para el espacio renglón. De todo esto, podemos concluir que los r renglones distintos de cero de una matriz escalonada U son linealmente independientes y también lo son las r columnas que contienen los pivotes distintos de cero.

La forma escalonada entonces se puede describir de la siguiente manera:

1. Primero vienen los renglones distintos de cero (de haber alguno se habría intercambiado).
2. Debajo de cada pivote, hay una columna de ceros que se obtiene por eliminación.
3. Cada pivote está a la derecha del pivote del renglón anterior, esto produce la figura escalonada.

Nos interesa localizar las soluciones. Si existe $Ax=b$, donde A es una matriz de $m \times n$, $x \in R^n$ y $b \in R^m$, las columnas de A las podemos ver como vectores de R^m, esto es:

$$A = (a_1, a_2, \cdots, a_n) \text{ con } a_i \in R^m$$

Donde, Ax podemos verla como una combinación lineal de los vectores columna de A, esto es:

$$Ax = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

Tal que $Ax=b$ tiene solución, si el vector Rm puede expresarse como combinación lineal de los vectores columna de A, de donde:

1. Si $m>n$, entonces los vectores columna son vectores de R^m y el número de vectores columna de A es menor que m, de donde el espacio generado por las columnas de A es un subespacio de R^m (de dimensión menor o igual a n).

   Por ejemplo:

   $3x_1 + 2x_2 = 3$
   $5x_1 + 6x_2 = 1$
   $3x_1 - 4x_2 = 5$

   Tenemos entonces que (3, 5, 3) y (2, 6, 4) son vectores en R^3, donde no pueden generar a R^3 pero si un subespacio de R^3. Si son linealmente independientes pueden generar un plano en R^3 y si no, una recta.

Tenemos que considerar además que:

a) El sistema no tenga solución; lo que sucede si b no se encuentra en el subespacio generado por las columnas de A.
b) Que tenga solución única; si los vectores columna de A son linealmente independientes y formen por lo tanto una base del espacio generado por las columnas de A.
c) Que tenga una infinidad de soluciones, esto ocurre si la dimensión del espacio generado por las columnas es menor que *n*, por lo cual, los vectores columna son linealmente dependientes y conforman un conjunto generador (pero no base, ya que generan un subespacio de dimensión menor que *n*) de donde, si *b* está en el espacio generado por las columnas de *A*, puede representarse en una infinidad de formas de combinaciones lineales de las columnas de *A*.

2. Si $m<n$ los vectores columna de *A* no pueden ser una base para R^m, ya que son más de los que se necesitan y para R^n solo pueden ser base para un subespacio.

Ejemplo 6.3

Comenzamos con el caso homogéneo $Ax=0$, donde *A* es la matriz:

$$A=\begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ 1 & -3 & 3 & 0 \end{bmatrix}$$

Que se puede reducir a $Ux=0$ de la siguiente manera:

$$Ux=\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Las incógnitas *x, y, z, w* van en dos grupos. Un grupo consta de las variables básicas, aquellas que corresponden a las columnas con pivotes distintos de cero (en este caso la 1a. y la 3a.) así que *x, z* son **variables básicas**. El otro grupo consta de las **variables libres**, que corresponden a las columnas sin pivotes, estas son la 2a y la 4a. columnas, entonces, *y, w* son las variables libres.

Para encontrar la solución general a $Ux=0$ ($Ax=0$) asignamos valores arbitrarios a las variables libres; esto es, y, w, entonces:

$$3z + w = 0 \rightarrow z = -1/3\, w$$
$$x + 3y + 3z + 2w = 0 \rightarrow x = -3y - w$$

Hay una infinidad de soluciones al sistema con dos parámetros libres w, y. La solución general es la combinación de:

$$x = \begin{bmatrix} -3y - w \\ y \\ -\tfrac{1}{3}w \\ w \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ -\tfrac{1}{3} \\ 1 \end{bmatrix}$$

Si damos valores $y=1$, $w=0$, tenemos la solución (−3, 1, 0, 0), si $y=0$, $w=1$ tenemos (−1, 0, −1/3, 1). Todas las soluciones son combinación lineal de estas dos.

Entonces las soluciones de $Ax=0$ forman un subespacio bidimensional, **el espacio nulo de *A***, del espacio de 4 dimensiones de todos los vectores x. Este espacio nulo lo podemos describir como un "plano" generado por (−3, 1, 0, 0) y (−1, 0, −1/3, 1). Las combinaciones de estos vectores forman un conjunto cerrado bajo la suma y el producto por un escalar y todas estas combinaciones constituyen el espacio nulo.

En el caso de $n>m$, no debe haber más de m pivotes distintos de cero, entonces debe haber al menos $n-m$ variables libres por lo cual: Cada sistema homogéneo $Ax=0$ tiene una solución no trivial, si tiene más incógnitas que ecuaciones ($n>m$): existe alguna solución x diferente de la trivial $x=0$.

Para el caso no homogéneo $b\neq 0$ tenemos $Ax=b$, aplicando operaciones elementales tenemos $Ux=c$:

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b \end{bmatrix}$$

No está claro que el sistema tenga solución. Ya que el tercer renglón de la matriz es igual a cero entonces:

$$b_3 - 2b_2 + 5b_1 = 0$$

El conjunto de vectores obtenibles b no es todo el espacio tridimensional. Es posible resolver $Ax=b$, si b está en el espacio columna de A generado por las columnas:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix} \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix}$$

Aunque son 4 vectores, solo obtenemos un plano en el espacio tridimensional (ya que la 2a columna es 3 veces la primera, y la cuarta es la primera más una fracción de la tercera). Son estas columnas las que no tienen pivotes.

Si suponemos que b está en el plano generado por el espacio columna, por ejemplo: $b=(1, 5, 5)$ efectuando eliminación y fijando las variables libres obtenemos:

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$

$3z + w = 3 \quad \rightarrow \quad z = 1 - ⅓w$

$x + 3y + 3z + 2w = 1 \quad \rightarrow \quad x = -2 - 3y - w$

De donde:

$$x = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ -⅓ \\ 1 \end{bmatrix}$$

Si comparamos la solución con la del caso homogéneo $Ax=0$, la única diferencia es la inclusión del vector $(-2, 0, 1, 0)$ que es una solución de las ecuaciones dadas, es una solución particular de $Ax=b$, y la solución general x es una suma de esta solución particular y la solución general de $Ax=0$. Geométricamente las soluciones se encuentran de nuevo en un plano en el espacio de 4 dimensiones, pero ahora no constituyen un subespacio, ya que el plano no pasa por el origen, el origen $x=0$ no es una solución cuando $b \neq 0$. El plano es paralelo al espacio nulo que se tenía antes, pero desplazado por el vector que da la solución particular.

Resumiendo, tenemos:

El propósito original de la eliminación, era simplificar un sistema de ecuaciones lineales sin alterar ninguna de las soluciones. Se reduce el sistema $Ax=0$ a $Ux=0$ y éste proceso es reversible. Por lo tanto, el espacio nulo de A es el mismo que el espacio nulo de U. De las m aparentes restricciones impuestas por la m ecuaciones $Ax=0$, solo r son independientes, están especificadas por cualesquiera r renglones de U distintos de cero.

> ***Definición 6.3***
> El espacio nulo de A: N(A) tiene dimensión $n-r$ y consta de todos aquellos vectores x para los cuales el sistema homogéneo $Ax=0$ (o su equivalente $Ux=0$) tiene solución. Se obtiene una base para el espacio nulo reduciendo el sistema original al sistema $Ux=0$, que tiene $n-r$ variables libres, correspondientes a las columnas de U que no contienen pivotes. Al espacio nulo también se le llama el núcleo de A y su dimensión $n-r$ nulidad de A.

De todo lo expuesto hasta aquí, podemos concluir que A no tiene el mismo espacio columna que U. La eliminación no altera el espacio de los renglones ni el espacio nulo, pero las columnas son completamente diferentes. Sin embargo, si el conjunto de columnas de A es independiente, entonces lo mismo es cierto para las correspondientes columnas de U y viceversa; entonces U nos puede ser útil en términos de indicarnos cuales son las columnas de A que forman una base.

Para encontrar una base de $R(A)$ consideremos el siguiente resultado:

Teorema 6.3
La dimensión del espacio columna $R(A)$ es igual al rango r, que también es igual a la dimensión del espacio fila: el número de columnas independientes es igual al número de renglones independientes. Una base de $R(A)$ está formada por aquellas r columnas de A correspondientes en U, a las columnas que contienen los pivotes distintos de cero.

Este teorema también afirma para el caso de matrices cuadradas, que si los renglones de una matriz cuadrada son l.i. también lo son las columnas.

Definición 6.4
El espacio nulo de A^t o espacio nulo izquierdo $N(A^t)$, es un subespacio de R^m que consta de todos aquellos vectores; y tales que $A^t\, y = 0$ o $y^t\, A = 0$, se le llama a y el vector nulo izquierdo de A.

Es fácil encontrar la dimensión de $N(A^t)$. Sabiendo que:

rango + nulidad = dimensión del espacio columna + dimensión del espacio nulo = número de columnas

Esta regla también se aplica para A^t la cual tiene m columnas.

Pero el rango de los renglones = rango de las columnas = r, entonces:

$$r + dim\, N(A^t) = m$$
$$dim\, N(A^t) = m - r$$

Esquemáticamente se puede ver en la figura 6.2.

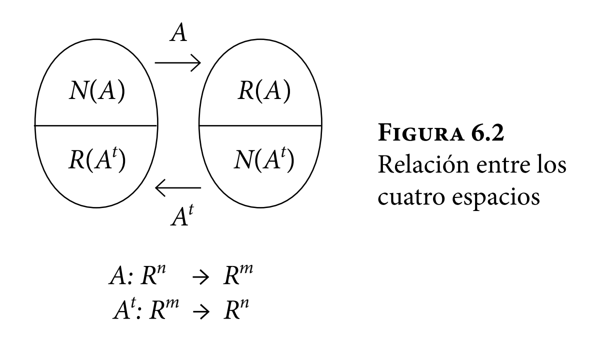


FIGURA 6.2
Relación entre los cuatro espacios

Sabemos ahora cuál es la dimensión de los cuatro espacios asociados a una matriz y resumimos este resultado en el siguiente teorema.

Teorema 6.4: Teorema fundamental del álgebra lineal (1a. Parte)

1. $R(A^t)$ = espacio renglón de A: dimensión = r
2. $N(A)$= espacio nulo de A: dimensión = $n-r$
3. $R(A)$ = espacio columna de A: dimensión = r
4. $N(A^t)$ = espacio nulo izquierdo de A: dimensión = $m-r$

Tal que:

$n = dim\ R(A) + dim\ N(A)$

$m = dim\ R(A^t) + dim\ N(A^t)$

Este teorema se ilustra con el siguiente ejemplo:

Ejemplo 6.4

Encuentre la dimensión y construya una base de los cuatros subespacios asociados a la matriz:

Solución

1. El espacio de los renglones de A: $R(A^t)$

   Escalonando la matriz A tenemos:

   $$\begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

   Entonces el vector (0,1,4,0) forma una base de los renglones de A y:

   $$Dim\left[R(A^t)\right] = 1$$

2. El espacio de las columnas. Buscamos las columnas cuyos pivotes sean distintos de cero y vemos que la única es la segunda columna $(1,2)^t$, de donde:

   $Dim\left(R(A)\right) = 1$ y una de sus bases $(1,2)^t$

3. Espacio nulo de A. $N(A)$ con dimensión $n-r \Rightarrow 4-1=3$ y para determinar una base, procedemos a resolver el sistema $Ax=0$, (también puede hacerse con $x=0$).

   De donde:

   $x_2 + 4x_3 = 0 \rightarrow x_2 = -4x_3$

   Sea $x_3 = t$, entonces $x_2 = -4t$ y haciendo $x_1 = u$ y $x_4 = v$, se obtiene:

   $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} u \\ -4t \\ t \\ v \end{bmatrix} = t\begin{bmatrix} 0 \\ -4 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Por lo cual:

$$\begin{bmatrix} 0 \\ -4 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Forman una base para $N(A)$, y $dim\big(N(A)\big) = n-r= 4-1= 3$, que son el número de vectores en la base.

4. Espacio nulo izquierdo de A :$N(A^t)$ con dimensión $m-r \to 2-1= 1$. Para encontrar una base procedemos a resolver $A^t\, y = 0$.

$$A^t y = \begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 4 & 8 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

De donde:

$y_1 = -2y_2$ haciendo $y_2 = t$

$y_1 = -2t$

Por lo cual una base de $N(A^t)$ está dada por:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

y: $dim\,(N(A^t)) = m - r = 2 - 1 = 1$

Ejemplo 6.5

Encuentre la dimensión y construya una base de los cuatro subespacios asociados a la matriz:

$$A= \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

Solución

1. El espacio de los renglones de A: $R(A^t)$:

$$A=\begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

Entonces los vectores $\{(1, 3, 3, 2), (0, 0, 3, 1)\}$ forman una base de los renglones de A y $Dim\ R(A^t) = 2$

2. El espacio de las columnas. Buscamos las columnas cuyos pivotes sean distintos de cero:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix}$$

$Dim\ R(A)= 2$

3. Espacio nulo de A: $N(A)$ con dimensión $n–r \rightarrow 4–2=2$ y para determinar una base, procedemos a resolver el sistema $Ax=0$, (también puede hacerse con $Ux=0$).

$$Ax=\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

De donde: $3x_3 + x_4 = 0 \rightarrow x_3 = -⅓x_4$

Sea $x_4=t$, entonces $x_3=-⅓t$ y haciendo $x_1 = -3x_2 - 3x_3 - 2x_4$ y $x_2 = u$ se obtiene:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3u - t \\ u \\ -t/3 \\ t \end{bmatrix} = u \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ -⅓ \\ 1 \end{bmatrix}$$

Por lo cual:

$$\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ -⅓ \\ 0 \end{bmatrix}$$

Forman una base para $N(A)$.

4. Espacio nulo izquierdo de A:$N(A^t)$ con dimensión $m-r \rightarrow 3-2 = 1$. Para encontrar una base procedemos a resolver $A^t\, y=0$.

$$A^t\, y = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \\ 3 & 9 & 3 \\ 2 & 5 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Por lo cual una base de $N(A^t)$ está dada por:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = t \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$$

Hemos visto hasta aquí que para encontrar una base del espacio renglón y una base para el espacio columna se procede a escalonar la matriz A en una matriz U, es entonces factible preguntar si estas matrices están relacionadas como antes, mediante una matriz triangular inferior L, tal que $A=LU$, la respuesta es sí existe una matriz L tal que $A=LU$, por ejemplo:

Si se tienen las siguientes matrices A y su forma escalonada U:

$$A= \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

$$U= \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Entonces la matriz L está dada por:

$$L=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$$

Y se verifica que $A=LU$, note además que L no es rectangular, sino cuadrada. Es una matriz del mismo orden $m=3$ que el número de renglones en A y en U. En este caso como se requirió de un intercambio de renglones entonces se introduce una matriz de permutación P con lo que enunciamos el siguiente teorema:

Teorema 6.5
A cada matriz A de orden $m\times n$ corresponde una matriz de permutación P, una matriz triangular inferior L con diagonal unitaria y una matriz escalonada U de orden $m\times n$ tales que $PA=LU$.

MATRICES DE RANGO UNO

Un ejemplo de una matriz de rango uno, $r=1$ está dado por la siguiente matriz:

$$A=\begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix}$$

Cada renglón es un múltiplo del primero de modo que el espacio renglón es unidimensional. Se puede escribir toda la matriz de la siguiente manera:

$$A=\begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix}=\begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix}(2,1,1)$$

Como el producto de un vector columna con un vector fila, además, todas las columnas son múltiplos del mismo vector el espacio columna tiene dimensión $r=1$.

Cualquier matriz de rango uno se puede factorizar en la forma simple $A=uv^t$. Todos los renglones son múltiplos del mismo vector v^t y todas las columnas son múltiplos de un mismo vector u.

6.2. SOLUCIÓN DE *M* ECUACIONES EN *N* INCÓGNITAS

Ya conocemos el proceso de eliminación para matrices cuadradas. La eliminación misma se realiza directamente sin mayor cambio, pero hay algunas diferencias al obtener la solución mediante la sustitución regresiva.

Consideremos la ecuación escalar $Ax=b$ (un sistema de una ecuación con una incógnita), podemos considerar las siguientes tres alternativas:

1. Si $a\neq 0$ entonces existe la solución $x=b/a$ para cualquier b y esta solución es única. Este es el caso **no singular**, de una matriz de 1×1 invertible.
2. Si $a=0$ y $b=0$ hay una infinidad de soluciones, cualquier x satisface $0x=0$. Este es el caso **indeterminado**, existe solución, pero no es única.
3. Si $a=0$ y $b\neq 0$ no hay solución a $0x=b$ este es el caso **inconsistente**.

Estas alternativas pueden ocurrir para matrices cuadradas. Con una matriz rectangular no podemos tener existencia y también unicidad para cada b. Esto se puede ver a través de la existencia de la inversa.

Sabemos que una matriz A tiene inversa si existen $BA=I$ y $AC=I$ donde $B=C=A^{-1}$. Ahora, del rango de una matriz es fácil decir que matrices tienen realmente estas inversas, se puede decir en general que existe una inversa solo cuando el rango es lo más grande posible.

El rango siempre satisface $r\leq m$ y $r\leq n$ ya que una matriz de $m\times n$ no puede tener más de m filas independientes *o* de n columnas independientes. Entonces si $r=m$ existe una inversa derecha y si $r=n$ existe una inversa izquierda.

> ***Definición 6.5***
> Sea A matriz de orden $m\times n$. Una matriz G de $n\times m$ (si existe) tal que $GA=I_n$ se dice inversa izquierda de A. Asimismo una matriz H de orden $n\times m$ (si existe) tal que $AH=I_m$ se dice inversa derecha.

Ejemplo 6.6

Sea la matriz de 2×1 siguiente:

$$A=\begin{bmatrix}1\\1\end{bmatrix}$$

Calcule (si existe) la inversa izquierda y derecha de A:

Solución:

Sea $G=[g_1, g_2]$ si $GA=1$:

$$[g_1, g_2]\begin{bmatrix}1\\1\end{bmatrix}=g_1+g_2=1$$

Si $g_1=\alpha$ entonces $g_2=1-\alpha$ de donde $G=[\,a, 1-\alpha]$ para $\alpha\in R$:

Sea $H=[h_1, h_2]$ tal que $AH=I$:

$$\begin{bmatrix}1\\1\end{bmatrix}(h_1, h_2)=\begin{bmatrix}h_1 & h_2\\h_1 & h_2\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

Pero $h_1=1$ y $h_1=0$ por lo que se afirma que no existe matriz derecha además el rango de $A=1<2=m$.

Teorema 6.6

Existencia: El sistema $Ax=b$ tiene al menos una solución x para toda b si y solo si las columnas de A generan R^m, entonces $r=m$. En este caso existe una inversa derecha H de $n\times m$ tal que $AH=I_m$, la matriz identidad de orden m. Esto es posible solo sí $m\leq n$.

UNICIDAD: El sistema $Ax=b$ tiene a lo sumo una solución x para cada b si y solo si las columnas son linealmente independientes, entonces $r=n$. En este caso existe una inversa izquierda G de $n\times m$ tal que $GA=I_n$, la matriz identidad de orden n. Esto es posible solo si $m\geq n$.

En el primer caso, una posible solución es $x=Hb$, ya que entonces $Ax=AHb=b$, pero habrá otras soluciones si hay otras inversas derechas.

En el segundo caso, si existe una solución a $Ax=b$, tiene que ser $x=GAx=Gb$, pero no puede haber otra solución.

Ejemplo 6.7

Considere la siguiente matriz A de 2×3 y de rango 2:

$$A=\begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}$$

Verifique si existen inversa izquierda o derecha de A.

Solución

Como $r=m=2$ el teorema garantiza una inversa derecha H:

$$AH=\begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}\begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

De hecho, hay muchas inversas derechas. Ya que el último renglón de H es totalmente arbitrario, entonces si escribimos un sistema $Ax=b$ con esta matriz A se puede afirmar que existe solución.

Para ver si existe inversa izquierda solo hay que verificar que $r=n$, en este caso $r=2<3$ por lo que no existe inversa izquierda.

Teorema 6.7

Una matriz A de $n\times n$ es invertible si y solo si rango $(A)=n$.

Ejemplo 6.8

Determine si la matriz siguiente es invertible hallando su rango.

$$A=\begin{bmatrix} 4 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}$$

Solución

Escalonando A, se tiene:

$$\begin{bmatrix} 4 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 3 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 1 & -12 \end{bmatrix}\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & -14 \end{bmatrix}$$

Por lo que rango $A=3$ y A es invertible.

6.3. SUBESPACIOS ORTOGONALES

Los espacios ortogonales juegan un papel fundamental en los espacios asociados a una matriz. Sabemos que en R^3 los subespacios se representan como rectas o planos que pasan por el origen, sin embargo, una recta puede ser ortogonal a otra o a un plano, pero dos planos no son ortogonales. (Strang, 1982).

Recordemos que la longitud de un vector $\|x\|$ en R^n está dada por:

$$\|x\| = \sqrt{x_1{}^2 + x_2{}^2 + x_3{}^2 + \cdots + x_n{}^2}$$

Geométricamente esto se obtiene aplicando la fórmula de Pitágoras n–1 veces, añadiendo en cada paso una dimensión más.

Si tenemos dos vectores x, y ¿Cómo podemos saber si son ortogonales?

En el caso de R^2 la respuesta se puede obtener usando trigonometría, en este caso x es ortogonal a y si forman un triángulo rectángulo, usando Pitágoras se tiene:

$$\|x\|^2 + \|y\|^2 = \|x-y\|^2$$

Para R^n se tiene:

$$(x_1{}^2+x_2{}^2+\cdots+x_n{}^2) + (y_1^2+y_2^2+\cdots+y_n{}^2) = (x_1-y_1)^2+\cdots+(x_n-y_n)^2$$

Desarrollando el lado derecho se tiene:

$$(x_1{}^2+x_2^2+\cdots+x_n{}^2) - 2(x_1y_1+\cdots+x_ny_n) + (y_1^2+y_2^2+\cdots+y_n^2)$$

Así, la igualdad se cumple si los "términos cruzados" son iguales a cero:

$$x_1y_1+\cdots+x_ny_n = 0$$

Esta cantidad no es otra, que el producto interior de x^ty en R^n; por lo que se puede afirmar que es igual a cero, solo si "x" e "y" son ortogonales.

Existe una relación sencilla entre independencia y ortogonalidad: si los vectores son mutuamente ortogonales, entonces son linealmente independientes.

> ***Definición 6.6***
> Dos subespacios V y W del mismo espacio R^n son ortogonales si cada vector $v\in V$ es ortogonal a cada vector $w\in W$: $v^tw=0$ para toda v y w.

Ejemplo 6.9

Suponga que V es el plano generado por $v_1= (1,0,0,0)$ y $v_2= (1,1,0,0)$ y W es una recta generada por $w_1= (0,0,4,5)$, como $w_1^t\, v_1=0$ y $w_2^t\, v_2=0$ entonces la recta W será ortogonal a todo el plano V.

En el caso de los cuatro subespacios fundamentales asociados a una matriz dos son subespacios de R^n: el espacio nulo y el espacio renglón o fila y los otros dos están en R^m, como se muestra en la figura 6.3

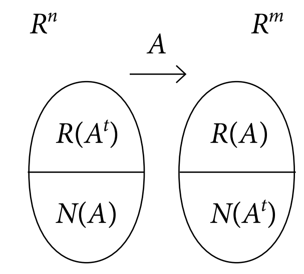


Figura 6.3 Los cuatro espacios como subespacios de R^n y R^m

Teorema 6.8

Para cualquier matriz A de orden $m\times n$, el espacio nulo y el espacio fila son subespacios ortogonales de R^n. Análogamente, el espacio nulo izquierdo y el espacio columna son subespacios ortogonales de R^m.

Demostración

Supongamos que $w\in N(A)$ y $v\in R(A^t)$, entonces $Aw = 0$ y v es de la forma $v=A^tx$ para algún vector x (esto es, v es una combinación de las columnas de A^t). Por lo tanto:

$$w^t v = w^t(A^t x) = (w^t A^t)x = (Aw)^t x = 0^t x = 0$$

Ejemplo 6.10

Sean la matriz A y su correspondiente U como sigue:

$$A=\begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \qquad U=\begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La segunda columna es básica y las otras tres son variables libres, por lo tanto, resolviendo para $Ux=0$, se obtiene que una base para $N(A)$ está dada por:

$$\left\{[1,0,0,0]^t, [0,-4,1,0]^t,\ [0,0,0,1]^t\right\}$$

La cual debe ser ortogonal a los renglones de A.

El espacio columna de A es unidimensional y está generado por la única columna básica $[1,2]^t$. Por otro lado, hallamos una base del espacio nulo izquierdo, dada por el vector $y^t = [-2,1]$, el cual es ortogonal al espacio columna:

$$[-2 \quad 1]\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0$$

Sin embargo, cabe aclarar que el espacio nulo $N(A)$ no solo contiene algunos de los vectores ortogonales al espacio fila, contiene todos esos vectores. El espacio nulo se formó con todas las soluciones de $Ax=0$.

> ***Definición 6.7***
> Dado un subespacio V de R^n, el espacio de todos los vectores ortogonales a V es el complemento ortogonal de V y se denota por $V^{\perp}$.

Así se tiene entonces que el espacio nulo $N(A)$ es el complemento ortogonal de $R(A^t)$: $N(A) = \left(R(A^t)\right)^{\perp}$. Al mismo tiempo es válida la relación opuesta: el espacio fila $R(A^t)$ contiene todos los vectores que son ortogonales al espacio nulo.

Al aplicar el mismo razonamiento a A^t, se produce el resultado dual: el espacio nulo izquierdo $N(A^t)$ y el espacio columna $R(A)$ son complementos ortogonales entre sí en R^m. Esto completa la segunda parte del teorema fundamental del álgebra lineal.

Teorema 6.9: Teorema fundamental del álgebra lineal (Segunda Parte)

$$N(A) = \left(R(A^t)\right)^{\perp}, \qquad R(A^t) = N(A)^{\perp}$$
$$N(A^t) = \left(R(A)\right)^{\perp}, \qquad R(A) = \left(N(A^t)\right)^{\perp}$$

La última igualdad significa que $Ax=b$ tiene una solución si y solo si b es ortogonal a $N(A^t)$: b está en el espacio columna si y solo si es ortogonal a cada solución y de la ecuación homogénea transpuesta $A^ty=0$.

Sin embargo, dos subespacios V y W pueden ser ortogonales sin ser complementos ortogonales entre sí. En R^3, la recta V generada por (1,0,0) es ortogonal a la recta W generada por (0,0,1), pero V no es igual $W^{\perp}$. El complemento ortogonal de W es un subespacio bidimensional (un plano) que contiene todos los vectores de la forma $(x_1, x_2, 0)$, esto es todos los generados por los vectores (1,0,0) y (0,1,0). La recta V solo puede ser parte de $W^{\perp}$ porque su dimensión es muy pequeña. Sin embargo, si las dimensiones son correctas, los dos subespacios ortogonales son necesariamente complementos ortogonales, así fue en el caso del espacio fila y el espacio nulo. Además, si $W=V^{\perp}$, esto asegura que las dimensiones son las correctas y automáticamente $V=W^{\perp}$. Simplemente se descompone el espacio en dos partes perpendiculares V y W.

Teorema 6.10

Si V y W son subespacios de R^n, entonces cualquiera de las siguientes condiciones los obliga a ser complementos ortogonales entre sí:

1. $W = V^{\perp}$ (W consta de todos los vectores ortogonales a V);
2. $V = W^{\perp}$ (V consta de todos los vectores ortogonales a W);
3. V y W son ortogonales y $Dim\ V + Dim\ W=n$.

Suponiendo cualquiera de estas tres condiciones equivalentes, cada vector x puede descomponerse de una sola manera en una suma $x=v+w$ con $v \in V$ y $w \in W$. Estas componentes, las proyecciones de x en V y W son ortogonales $v^tw=0$.

Una x arbitraria se descompone en x_r+x_n, y A transforma la componente x_r del espacio fila en un vector $Ax_r=Ax$ en el espacio columna, mientras que transforma la componente x_n del espacio nulo en cero.

Se muestra este teorema en la siguiente figura 6.4

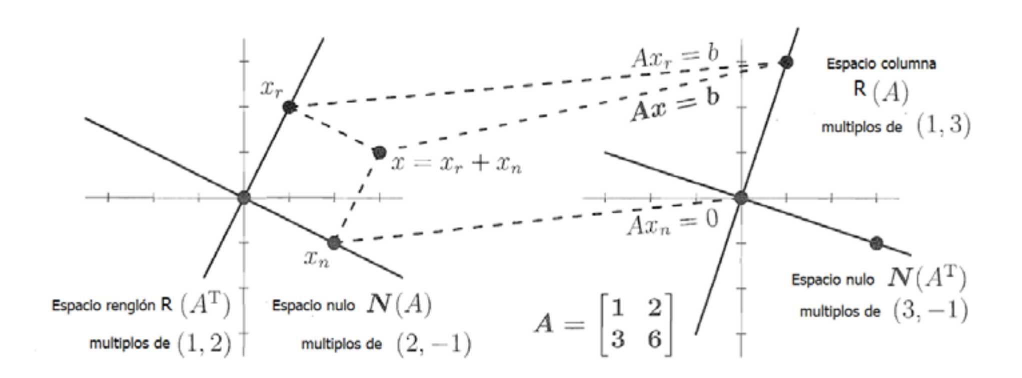


Figura 6.4 Teorema de álgebra lineal segunda parte (con ejemplo). Fuente Strang (2007)

Ejemplo 6.11

Sea la matriz A como sigue:

$$A=\begin{bmatrix} 0 & 1 & 4 \\ 0 & 2 & 8 \end{bmatrix}$$

Su correspondiente forma escalonada queda como sigue:

$$U=\begin{bmatrix} 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

De aquí que: $Dim\ R(A^t) = 1,\ Dim\ N(A) = 2$

Una base de $R(A^t) = \{(0,\ 1,\ 4)\}$,

Una base de $N(A)$ está dada por $\{(1,\ 0,\ 0),(0,\ 4,\ 1)\}$ entonces cualquier $x \in R^3$ se puede escribir como:

$$x=\left\{a_1\begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}\right\}+\left\{a_2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}+a_3\begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}\right\}$$

base de $R(A^t)$ + base de $N(A)$

6.4. ESPACIOS CON PRODUCTO INTERIOR Y EL PROCESO DE GRAM-SCHMIDT

En la sección anterior se consideraron subespacios ortogonales, y en este sentido al considerar una base se tiende a elegir una base ortogonal. Como explica Strang (1982): “Si la idea de base es uno de los pasos clave para conectar la geometría de un espacio vectorial con el álgebra entonces la especialización a una base ortogonal no queda muy atrás”. Lo que nos lleva al siguiente paso, si ya tenemos una base ortogonal, los podemos normalizar para que sean vectores unitarios, de esto trata esta sección.

> ***Definición 6.8***
> Un producto interior en un espacio vectorial *V* es una función que a cada par de vectores *x* y *y* en *V* asocia un número real$<x, y>$ de tal manera que se satisfacen los siguientes axiomas, para todos los vectores *x, y, z* en *V* y cualquier escala *k*.

1. $<x, y> = <y, x>$	axioma de la simetría
2. $<x+y, z> = <x, z> + <y, z>$	axioma de la aditividad
3. $<kx, y> = k <x, y>$	axioma de la homogeneidad
4. $<x, x> \geq 0 y <x, x> = 0 <=> x = 0$	axioma de positividad.

Un espacio vectorial que tiene definido un producto interior se denomina espacio vectorial con producto interior.

Las siguientes propiedades adicionales son una consecuencia inmediata de los cuatro axiomas del producto interior:

a) $<0, x> = <x, 0> = 0$
b) $<x, y+z> = <x, y> + <x, z>$
c) $<x, ky> = k<x, y>$.

Ejemplo 6.12

Considere el espacio R^n con producto interior dado como:

$$< x,y > = x^t y = \sum_{i=l}^{n} x_i y_i$$

Donde x, y son vectores columna en R^n.

Note que <., ,> o <./.> es un producto interior ya que:

1. $<x, y> = x^t y = y^t x = <y, x>$
2. $<x+y, z> = (x+y)^t z = (x^t+y^t)z = x^t z + y^t z = <x, z> + <y, z>$
3. $<kx, y> = (kx)^t y = kx^t y = k<x, y>$
4. $<x, x> = x^t x = \Sigma x_i^2 \geq 0$ *si* $\Sigma x_i^2 = 0$ entonces $x = 0 \in R^n$

Ejemplo 6.13

Sean $x = <x_1, x_2>$, $y = <y_1, y_2>$ vectores en R^2, entonces:
$<x, y> = 3x_1 y_1 + 2x_2 y_2$

Define un producto interior.

Solución

Para verificar esta afirmación observe que:

1. $<x, y> = 3x_1 y_1 + 2x_2 y_2 = 3y_1 x_1 + 2y_2 x_2 = <y, x>$
2. $<x+y, z> = 3(x_1+y_1)z_1 + 2(x_2+y_2)z_2 = (3x_1z_1+2x_2 z_2) + (3x_1z_1+2x_2z_2) =$
   $<x, z> + <y, z>$
3. $<kx, y> = 3(kx_1)y_1 + 2(kx_2)y_2 = k(3x_1y_1+2x_2y_2) = k<x, y>$
4. $<x, x> = 3x_1 x_1 + 2x_2 x_2 = 3x_1^2 + 2x_2^2 \geq 0$ y $3x_1^2 + 2x_2^2 = 0$ si y solo si $x_1=x_2=0$.

El producto interior de este ejemplo es distinto al producto interior de R^2 lo que muestra que un espacio vectorial puede tener más de un producto interior definido en él.

Ejemplo 6.14

Sean $x = <x_1, x_2>$, $y = <y_1, y_2>$ vectores en R^2, entonces:
$<x, y> = x_1 + y_1$

Verifique si $<x, y>$ define un producto interior.

Solución

Para verificar esta afirmación observe que:

1. $<x, y> = x_1 + y_1 = y_1 + x_1 = <y, x>$
2. $<x+y, z> = (x_1 + y_1) + z_1 = (x_1 + z_1) + y_1 = <x+z, y> \leq <x, z> + <y, z>$

Por lo que se puede concluir que $<x, y>$ así definido no es un producto interior.

Ejemplo 6.15

(Caso de polinomios) Sea P_n el espacio vectorial de los polinomios de grado n. Un producto interior en este espacio está dado por:

$$\langle x,y\rangle = \int_a^b x\,(t)\,y(t)\,dt$$

1. $<x,y> = \int_a^b x\,(t)\,y(t)\,dt = \int_a^b y(t)\,x(t)\,dt = <x,y>$

2. $<x,y + z> = \int_a^b x(t)\,\left[y(t) + z(t)\right]\,dt = \int_a^b \left[x(t)\,y(t) + x(t)\,z(t)\right]\,dt =$

   $\int_a^b x(t)\,y(t)\,dt + \int_a^b x(t)\,z(t)\,dt = <x,y> + <x,z>$

3. $<kx,y> = \int_a^b kx(t)\,y(t)dt = k\int_a^b x(t)\,y(t)dt = k<x,y>$

4. $<\mathrm{x},\mathrm{y}> = \int_a^b x^2\,(t)\,dt = 0$ si y solo si $x=0$

Definición 6.9

Si V es un espacio con producto interior, entonces se llama un espacio normado si existe una función $\|..\| : V \to R$ tal que:

$\| x + y \| \leq \| x \| + \| y \|$

$\| \alpha x \| = |\alpha| \, \| x \|$

$\| x \| \geq 0$ y $\| x \| = 0 \leftrightarrow x = 0$

Ejemplo 6.16

Muestre que el espacio $V=R^n$ es normado con cualquiera de las siguientes normas:

a) $\| x \|_1 = \sum_{i=l}^{n} | x_i |$

b) $\| x \|_1 = \left(\sum_{i=l}^{n} x_i^2 \right)^{½}$

c) $\| x \|_\infty = max \left\{ | x_i | \; i=1,2, \cdots n \right\}$

Solución

Vamos a demostrar que efectivamente a), b) y c) son normas, esto es:

a) $\| x+y \|_1 = \sum | x_i + y_i | \leq \sum | x_i | + \sum | y_i | = \| x \| \; 1 + \| y \|_1$

$\| \alpha x \| = \sum | ax_i | = | \alpha | \sum x_i = | \alpha | \; \| x \|_1$

$\| x \| = \sum | x_i | \geq 0 \; y \; \| x \| = 0 \leftrightarrow x = 0$

b) $\| x + y \|_2 = \sum (x_i + y_1)^2 = \sum (x_i^2 + 2x_i y_i + y_i^2) =$

$\sum x_i^2 + 2 \sum x_i \, y_i + \sum y_i^2 = \| x \|_1 + \| y \|_1$

c) $\| x + y \|_\infty = max \left\{ |x_i + y_i |, \; i=1,2, \cdots n \right\} \leq max \left\{ | x_i | + | y_i | \; i=1,2, \cdots n \right\}$

$max \left\{ \; | x_i | \; i= 1,2, \cdots, n \right\} + max \left\{ | y_i | \; i= 1, 2 \cdots n \right\} = \| x \|_\infty + \| y \|_\infty$

$\alpha x \|_\infty = max \left\{ |\alpha x_i | \; i= 1, 2, \cdots, n \right\} = max \left\{ |\alpha| \; x_i | \; i= 1,2,\cdots, n \right\} =$

$| \alpha | \; max \left\{ | x_i | \; i= 1, 2, \cdots, n \right\} = \alpha \| x \|_\infty$

6.5. ÁNGULO EN LOS ESPACIOS VECTORIALES, PROYECCIONES ORTOGONALES Y PROCESO DE GRAM-SCHMIDT

Ya vimos en la sección anterior que el hecho de que dos vectores x y y sean ortogonales se puede expresar como $x^t y = 0$ donde x, y son vectores columna.

Considere el espacio vectorial R^2 y los vectores a y b donde sus longitudes están dadas por a y b respectivamente. Sea α el ángulo que forma el vector a con el eje de las x y β el ángulo del vector b con el mismo eje como se ilustra en la figura 6.5.

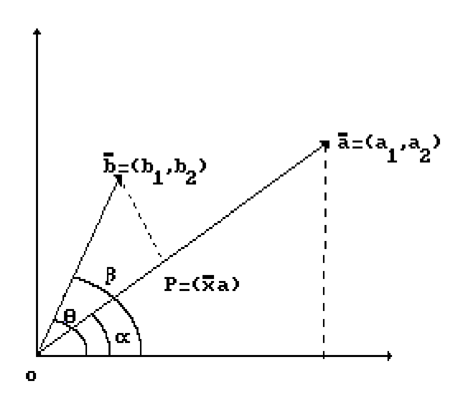


Figura 6.5 Ángulo en los espacios vectoriales

Entonces se definen el seno y coseno de *a* como:

$$sen\ \alpha = \frac{a_2}{\|a\|} \qquad cos\ \alpha = \frac{a_1}{\|a\|}$$

Y para β se tiene:

$$sen\ \beta = \frac{b^2}{\|a\|} \qquad cos\ \beta = \frac{b_1}{\|a\|}$$

Si $\Theta = \beta - \alpha$ entonces usando una identidad trigonométrica se tiene:

$$cos\ \Theta = cos\ \beta\ cos\ \alpha + sen\ \beta\ sen\ \alpha = a_1\ b_1$$

Por lo tanto, el coseno del ángulo entre dos vectores cualesquiera *a* y *b* está dado por:

$$cos\ \Theta = \frac{\langle \alpha, b \rangle}{\|\alpha\|\ \|b\|}$$

Si queremos encontrar la proyección de un punto del vector *b*, sobre el vector *a*, este punto debe ser algún múltiplo $p=xa$ del vector *a* y el problema se reduce a calcular el coeficiente *x*. Todo lo que necesitamos es el hecho de que la recta desde *b* hasta el punto más cercano $p=xa$ es ortogonal al vector *a*, como se puede ver en la figura 6.4, entonces:

$$(b-xa)\perp a \quad \text{o} \quad <a, b-xa> = 0$$

De donde:

$$x = \frac{< a, b >}{< a, a>} = \frac{< a, b >}{\| a \|^2}$$

Entonces la proyección p del punto b sobre la recta generada por el vector a está dada por:

$$p = \frac{< a,b >}{\| a \|^2} a$$

De esta relación se puede encontrar nuevamente la desigualdad de Cauchy–Schwarz ya que la distancia al cuadrado del punto a la recta es:

$$\left\| b - \frac{\langle a,b \rangle}{\| a \|^2} a \right\| = \frac{\| b \|^2 \| a \|^2 - \langle a,b \rangle}{\| a \|^2}$$

Como la distancia de un punto a una recta es mayor o igual a cero lo mismo sucede con el cuadrado de la distancia entonces:

$\| b \|^2 \| a \|^2 - \langle a,b \rangle \geq 0$
$\| b \|^2 \| a \|^2 \geq \langle a,b \rangle$ que es la desigualdad de Schwarz

Al seleccionar una base se busca que ésta simplifique la solución del problema que se trate. En la mayoría de los casos, la mejor selección será la base en la cual todos los vectores son ortogonales entre sí, es así como definimos:

> ***Definición 6.10***
> Se dice que un conjunto de vectores en un espacio con producto interior es ortonormal; si cada vector del conjunto tiene norma 1 y si además dos vectores cualesquiera distintos en el conjunto son ortogonales.

Ejemplo 6.17
Sean:
$v_1 = (0, 1, 0)$, $v_2 = (1/\sqrt{2}, 0, 1/\sqrt{2})$, $v_3 = (1/\sqrt{2}, 0, -1/\sqrt{2})$, el conjunto $S = \{v_1, v_2, v_3\}$ es ortonormal si en R^3 se tiene que con el producto interior euclideano:

$$<v_1, v_2> = <v_2, v_3> = <v_1, v_3> = 0$$

Y, además:

$$\|V_1\| = \|V_2\| = \|V_3\| = 1$$

Ejemplo 6.18
Si v es un vector en un espacio con producto interior, y v es diferente de cero, entonces el vector:

$$\frac{1}{\|V\|}V{=}1 \text{ tiene norma } 1$$

Ya que aplicando $\| ku \| = \| k \| \| u \|$se tiene:

$$\left\| \frac{1v}{\| v \|} \right\| = \frac{1}{\| v \|} \| v \| = 1$$

Al proceso de multiplicar un vector v diferente de cero por el recíproco de su longitud para obtener un vector de norma 1 se denomina **normalización de v.**

Teorema 6.11 Teorema de la base ortonormal (Gram-Schmidt)
Todo subespacio distinto de cero en R^n tiene una base ortonormal.

Demostración
Queremos producir de dos vectores independientes a y b dos vectores perpendiculares v_1 y v_2. Claramente, el primero puede ir en la dirección de a $v_1{=}a$. El problema es entonces encontrar un segundo vector que sea perpendicular.

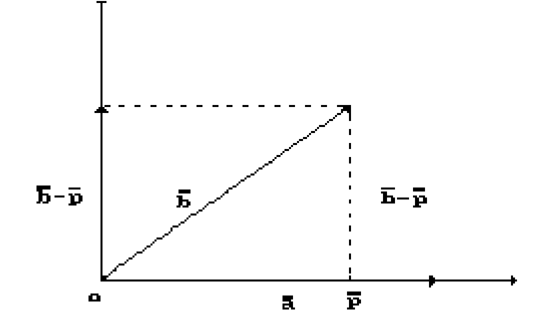


Figura 6.6
Construcción de vectores ortogonales

En la figura 6.6 se puede ver que ese vector es $b–p$ ya que se le sustrajo la proyección de b en la dirección de a, entonces el segundo vector queda:

$$v_2 = b - p = b - \frac{< a,b>}{\|a\|^2} a = b - \frac{<v_1, b>}{\|v_1\|^2} v_1$$

Si hay un tercer vector independiente entonces la idea es la misma, sustraemos las proyecciones de c en las dos direcciones v_1 y $v_{2.}$

$$v_3 = c - \frac{<v_1, c>}{\|V_1\|^2} v_1 - \frac{< v_2, c>}{\|V_2\|^2} v_2$$

Es inmediato que v_3 es perpendicular a v_1 y $v_{2.}$:

Para que estos vectores sean ortonormales deben transformarse en vectores unitarios entonces:

$$q_1 = \frac{v_1}{\|V_1\|} \quad q_2 = \frac{v_2}{\|V_2\|} \quad q_3 = \frac{v_3}{\|V_3\|}$$

Este proceso se puede resumir de la siguiente manera:

Es posible convertir cualquier conjunto de vectores independientes $a_1, a_2, \cdots, a_n$ en un conjunto de vectores ortogonales mediante el proceso de Gram–Schmidt: primero se fija $a_1 = v_1$ después cada v_i es ortogonal a las $v_1, \cdots, v_{i-1}$ precedentes:

$$v_1 = a_1 - \frac{< v_1, a_1 >}{\|V_1\|^2} v_1 - \ldots - \frac{< v_{i-1}, a_1 >}{\|V_{i-1}\|^2} v_2$$

Ejemplo 6.19

Aplique el proceso de Gram–Schmidt a los siguientes vectores independientes: $a_1 = [1, 1, 0]$, $a_2 = [1, 0, 1]$, $a_3 = [0, 1, 1]$.

Solución

Sea $a_1 = v_1$, v_2 y v_3 se calculan:

$v_2 = a_2 - \frac{1}{2} v_1 = [\frac{1}{2}, -\frac{1}{2}, 1]$

$v_3 = a_3 - \frac{1}{2} v_1 - \frac{1}{3} v_2 = [-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}]$

Los vectores ortonormales finales son:

$$q_1 = \sqrt{½}\,[1, 1, 0]$$
$$q_2 = \sqrt{⅔}\,[½, -½, 1]$$
$$q_3 = \sqrt{¾}\,[-⅔, ⅔, ⅔]$$

6.6. MATRICES DE PROYECCIÓN Y MÍNIMOS CUADRADOS

Considere ahora la proyección p de un vector sobre un subespacio general $S=\{a_1, a_2,\cdots,a_k\}$ de R^n donde las a_i son vectores independientes de R^n. Los detalles se elaboran para el caso $k=2$, pero los cálculos son igualmente válidos para el caso general. Recordemos que $\{a_1, a_2\}$ corresponde a un plano en R^n que contiene al origen, y sus miembros son todas las combinaciones lineales de a_1 y a_2.

La figura 6.7 muestra al vector proyección p de b en $\{a_1, a_2\}$.

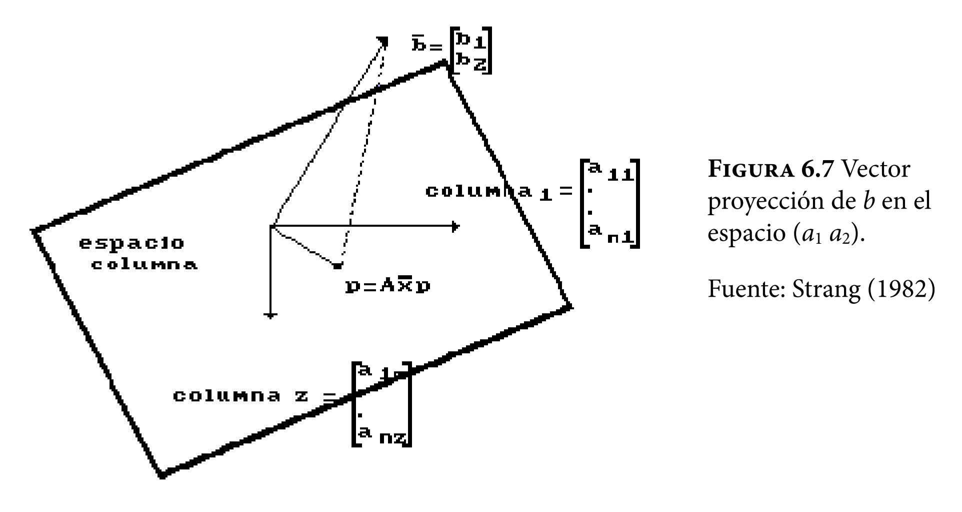


FIGURA 6.7 Vector proyección de b en el espacio (a_1 a_2).

Fuente: Strang (1982)

Observe que p satisface:

1. p debe estar en el subespacio $S = \{a_1, a_2\}$ y;
2. $b-p$ debe ser perpendicular a cada vector de S.

Si escribimos los vectores de R^n como vectores columna, el subespacio S es el espacio columna de la matriz A de $nx2$ cuyas columnas son a_1 y a_2 entonces como $Ax=b$ tiene solución si y solo si b está en el espacio columna de A, todos

los vectores de S tienen la forma Ax donde $x = [x_1, x_2]^t$. Como p está en el espacio S entonces $p=Ax_p$ donde $x_p = [r_1, r_2]^t$, r_1, r_2 escalares, como $b–Ax_p$ debe ser perpendicular a cada vector de S entonces el producto punto de $b–Ax_p$ y Ax debe ser igual a cero, para toda x.

$$<b-Ax_p, Ax> = 0$$

O bien: $(Ax)^t (b - Ax_p) = x^t(A^tb - A^tAx_p) = 0$

Es decir, el producto punto de los vectores x y $A^tb - A^tAx_p$ debe ser cero para todos los vectores x, pero esto solo sucede si:

$$A^t b - A^tA x_p = 0$$

La matriz A^tA de 2×2 es invertible* pues tiene el mismo rango que A, entonces despejando x_p se tiene:

$$x_p = (A^tA)^{-1} A^t b$$

Denotando el vector proyección $p=Ax_p$ de b sobre S por medio de $p=b_s$ y escribiendo b como un vector columna obtenemos la forma general:

Sea $S = \{a_1,\cdots,a_k\}$ un subespacio de R^n y sea la matriz A cuyas columnas son los vectores $a_1, a_2,\cdots, a_k$. El vector proyección de b en R^n sobre S está dado por:

$$b_s = A(A^tA)^{-1} A^tb$$

Este resultado se puede usar para una proyección de un vector en otro, por ejemplo:

Ejemplo 6.20

Sea el vector $a=(2, 4, 3)$, encuentre el vector proyección sobre a del vector $b=(1, 2, 3)$.

Solución

$$p=\frac{< a,b >}{\|a\|^2}a = \frac{19}{29}(2, 4, 3)$$

Considerando la fórmula $b_s = A(A_t A)^{-1} A^t b$ se tiene que la matriz A consiste de una sola columna que es el vector a, haciendo cálculos se tiene:

$$A^t A = 29$$

$$\begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} \frac{1}{29} (2,4,3) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{29} \begin{bmatrix} 4 & 8 & 6 & 3 \\ 8 & 16 & 12 & 2 \\ 6 & 12 & 9 & 3 \end{bmatrix}$$

$$b_s = A(A^t A)^{-1} A^t b =$$

$$= \frac{1}{29} \begin{bmatrix} 38 \\ 76 \\ 57 \end{bmatrix}$$

A la matriz $P_s = A(A^t A)^{-1} A^t$ se le conoce como la matriz de proyección para el subespacio S.

Ejemplo 6.21

Encuentre la matriz de proyección para el plano $x_2 x_3$ de R^3.

Solución

El plano $x_2 x_3$ es el subespacio S generado por los vectores $e_2\ e_3$ que son las columnas de la siguiente matriz A:

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Donde $A^t A = Iy$, por lo tanto:

$$P_s = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Así, P_s proyecta cada vector b en R^3 en el plano x_2x_3 de la siguiente manera:

$$P_s = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ b_2 \\ b_3 \end{bmatrix}$$

Ejemplo 6.22

Sean la matriz A y el vector b como sigue:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 4 \\ 3 \\ 9 \end{bmatrix}$$

Encuentre la proyección del vector b en el sobre el plano generado por los vectores columna de la matriz A.

Solución

El espacio columna de A es el plano x–y en el espacio tridimensional ya que ambas columnas terminan con un cero. La proyección de b sobre este plano no altera las componentes x, y que son 4 y 3 pero desaparecerá la componente z y se puede confirmar.

$$A^t A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 5 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 7 \\ 7 & 29 \end{bmatrix}$$

$$(A^t A)^{-1} = (1/9) \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix}$$

$$b_2 \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$$

Teorema 6.12

Para cualquier matriz A de $m \times n$, de rango r, la matriz simétrica A^tA de $n \times n$ también tiene rango r.

Demostración

Trabajando con los espacios nulos se tiene que si v es cualquier vector solución del sistema $Ax=0$, de modo que $Av=0$ entonces, al multiplicar a la izquierda ambos miembros de esta última ecuación por A^t, vemos que v también es una solución al sistema $(A^t)Ax=0$. Recíprocamente, supongamos que $(A^t)Aw=0$ para un vector w de $nx1$. Entonces:

$$0 = w^t \left[(A^t)Aw\right] = (Aw)^t(Aw)$$

Que se puede escribir como $[Aw]^2$. Esto es, Aw es un vector de magnitud cero, en consecuencia, $Aw = 0$. Resulta así que A *y* $(A^t)A$ tienen el mismo espacio nulo. Como tienen el mismo número de columnas, se concluye que el rango de A y el rango de $(A^t)A$ son iguales.

De este teorema se desprende el siguiente resultado: si A tiene columnas linealmente independientes de modo que $r=n$, entonces A^tA es una matriz cuadrada, simétrica e invertible.

MATRICES ORTOGONALES

Una matriz ortogonal es simplemente una matriz cuadrada con columnas ortonormales. Se usará la letra Q para denotar una matriz ortogonal y a $q_1,\cdots,q_n$ para denotar sus columnas.

Como Q es cuadrada de $n\times n$ entonces si sus columnas son independientes, esto es su rango r es tal que $r= n$ y es invertible, entonces: si Q^t es inversa izquierda entonces es la inversa y $QQ^t = I$.

Una matriz ortogonal tiene las siguientes propiedades:

$$Q^tQ = I = QQ^t$$
$$Q^t = Q^{-1}$$

Q no solo tiene las columnas ortonormales, sino que también sus filas son ortonormales, en otras palabras, si Q es ortogonal también lo es Q^t.

Además, cumple: La multiplicación por una matriz ortogonal Q preserva la longitud.

$$\| Qx \| = \| x \| \text{ para toda vector } x$$

Y también preserva los productos internos:

$(Qx)^t\,(Qy) = x^t$ y para todos los vectores x,y.

Ejemplo 6.23

Sea la matriz Q como sigue:

$$Q=\begin{bmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{bmatrix}$$

Su transpuesta es igual a su inversa:

$$Q^t = Q^{-1} = \begin{bmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{bmatrix}$$

Q rota cada vector en un ángulo Q y Q^t lo rota de regreso en $-Q$.

Ejemplo 6.24

Verifique que la matriz siguiente es una matriz ortogonal y encuentre su inversa.

$$A = (1/7)\begin{bmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \\ 6 & 2 & -3 \end{bmatrix}$$

Solución

$$A^t A = 1$$

$$A^{-1} = A^t = (1/7)\begin{bmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \\ 6 & 2 & -3 \end{bmatrix}$$

6.7. LA PSEUDOINVERSA Y LA DESCOMPOSICIÓN EN VALOR SINGULAR

En este punto cabe hacernos la siguiente pregunta: ¿Cuál es la solución óptima x para el sistema inconsistente $Ax=b$? deberíamos tener una regla que especifique para una x dada cualquier matriz A y cualquier lado derecho b.

Para cada b, Ax debe estar en el espacio de las columnas de A ya que es una combinación de las columnas ponderadas por las componentes de x. Por lo tanto, la elección óptima Ax es el punto p en el espacio columna que está más cerca a la b dada. En otras palabras, tenemos que proyectar a b sobre el espacio columna:

$$Ax=b_s = p$$

Claramente x está determinada cuando hay solo una combinación de las columnas de A que producen p, los pesos de esta combinación serán las componentes de x. Si se tiene solución única existen varias condiciones equivalentes:

a) Las columnas de A son linealmente independientes
b) El espacio nulo de A solo contiene al cero
c) El rango de A es n
d) La matriz cuadrada AtA es invertible.

Si estas condiciones no son válidas, entonces x no está determinada en forma única y:

$$x = (A^tA)^{-1} A^t b$$

Es la expresión de dicha solución. Si A es una matriz invertible entonces x coincide con la única solución al sistema original:

$$x = A^{-1} (A^t)^{-1} A^t b = A^{-1} b \qquad (6.1)$$

Pero si A no es invertible definimos la pseudoinversa $A+$ (o inversa de Moore–Penrose), por lo tanto, si A es invertible $A+ = A^{-1}$.

Cuando la matriz A cumple con las cuatro condiciones anteriores, la pseudoinversa es la inversa izquierda que aparece en (6.1):

$A+ = (A^tA)^{-1} A^t$

Pero la pseudoinversa queda por definirse cuando no son válidas las condiciones a), d) y x no está determinada en forma única por $Ax=p$. Tenemos que escoger uno de los muchos vectores que satisfacen

$Ax=p$ y la elección será la solución óptima $x=a+b$ al sistema inconsistente $Ax=b$.

Esta elección se efectúa de acuerdo a la regla siguiente: la solución óptima de entre todas las soluciones de $Ax=p$ es aquella con longitud mínima. La clave para encontrarla es recordar que el espacio fila y el espacio nulo de A son complementos ortogonales en R^n. Esto significa que cualquier vector puede descomponerse en dos piezas perpendiculares, su proyección sobre el espacio fila y su proyección sobre el espacio nulo. Suponiendo que aplicamos esta descomposición a una de las soluciones x_0 de la ecuación $Ax=p$. Entonces $x_0=x_r+x_n$ donde x_r está en el espacio fila y x_n está en el espacio nulo. Hay ahora tres puntos importantes:

1. La componente x_r es una solución de $Ax=p$ ya que $Ax_n=0$,

   $$Ax_0 = A(x_r + x_n) = Ax_r = p$$

2. Todas las soluciones de $Ax=p$ comparten ésta misma componente xr en el espacio fila y difieren solamente en la componente x_n en el espacio nulo. La solución general es la suma de una solución particular x_r y una solución arbitraria x_n de la ecuación homogénea.

3. La longitud de dicha solución $x_r + x_n$ obedece la ley de Pitágoras, ya que las dos componentes son ortogonales:

   $$\| x_r + x_n \|^2 = \| x_r \|^2 + \| x_n \|^2$$

De lo anterior se desprenden las siguientes conclusiones:

1. La solución que tiene longitud mínima es x_r. Deberíamos elegir como cero a la componente en el espacio nulo, dejando una solución situada completamente en el espacio fila.

2. La solución óptima en mínimos cuadrados para cualquier sistema $Ax=b$ es el vector x_r determinado por las dos condiciones:
   2.1 Ax es igual a la proyección de b sobre el espacio columna de A.
   2.2 x está en el espacio fila de A.
3. La matriz que resuelve $Ax=b$ es la pseudoinversa $A+$ definida por $x=A^{+}b$.

Una mejor manera de entender la pseudoinversa es viéndola geométricamente, como se muestra en la figura 6.8

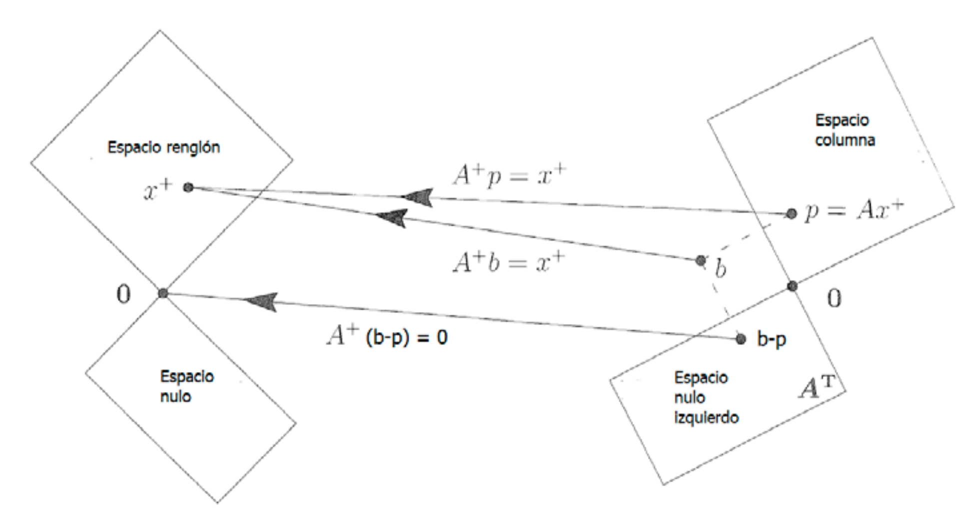


Figura 6.8 Representación gráfica de la pseudoinversa. Fuente: Strang(2007)

La matriz $A+$ combina el efecto de dos pasos separados: proyecta a b sobre el punto p y después encuentra el único vector x en el espacio fila que resuelve $Ax=p$. Un caso extremo es cuando b es perpendicular al espacio columna, en otras palabras, cuando b está en el espacio nulo izquierdo. Entonces $p=0$ *y* $x=0$, $A+$ envía todo $b–p$ a cero. En el otro extremo tenemos a b dentro del espacio columna entonces $p=b$ y encontramos x al "invertir" A (ya se había dicho que A es invertible si la consideramos como una aplicación de su espacio fila en su espacio columna con $A+$ como la inversa).

Una b arbitraria está entre estos dos extremos: la componente p se invierte para dar x y la otra componente $b–p$ se aniquila. De esta descripción y de la figura 6.7 podemos listar algunas propiedades básicas de la pseudoinversa:

1. $A+$ es una matriz de $n\times m$ comienza con el vector $b\in R^m$ y produce el vector $x\in R^n$.

2. El espacio columna de $A+$ es el espacio fila de A y el espacio fila de $A+$ es el espacio columna de A, de aquí que rango A = *rango* $A+$.
3. La pseudoinversa de $A+$ es A.
4. En general $AA+ \neq I$, ya que es posible que A no tenga inversa derecha, pero $AA+$ siempre es igual a la proyección P sobre el espacio columna:

$$AA^{+}b = Ax = p = Pb$$
$$AA^{+} = P$$

Ejemplo 6.25

> ***Condiciones de Penrose***
> Si $A \in R^{m \times n}$, entonces existe una única matriz $A^{+} \in R^{n \times m}$ que satisface las siguientes cuatro condiciones:
> 1. $AA^{+} A = A$
> 2. $A^{+} A A^{+} = A^{+}$
> 3. $A^{+} A = (A^{+}A)^{t}$
> 4. $AA^{+} = (AA^{+})^{t}$

Sea la matriz A una matriz no invertible:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Observe que esta no es otra que la matriz de proyección que manda cualquier punto de R^3 al plano $x–y$. De esta forma el espacio columna y el espacio fila de A coinciden con el plano $x–y$ en R^3 que contiene todos los vectores $(x, y, 0)$. El espacio nulo es el eje z, que es ortogonal al espacio fila. Para encontrar x se proyecta b sobre el espacio columna:

Si:

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \text{ entonces } p = pb = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

Y resolviendo $Ax=p$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

Se tiene:

$x_1 = b_1 \qquad x_2 = b_2 \qquad y \qquad x_3$ es arbitraria

Se elige entre esta familia infinita de soluciones aquella que tenga longitud mínima, donde la tercer componente debe ser cero, esto deja a $x=(b_1, b_2, 0)^t$ que está en el espacio fila (el plano x–y) y la pseudoinversa A^+ está dada por A misma. Esto sucede porque A actúa como la matriz identidad en la manera como aplica su espacio fila en su espacio columna y la pseudoinversa ignora todo lo demás. Resumiendo, se tiene que la solución óptima de $Ax=b$ que es un conjunto inconsistente:

$$x_1 + 0x_2 + 0x_3 = b_1$$
$$0x_1 + x_2 + 0x_3 = b_2$$
$$0x_1 + 0x_2 + 0x_3 = b_3$$

Consiste en satisfacer las dos primeras ecuaciones y hacer $x_3=0$.

6.8. RESUMEN

1. La dimensión de los cuatro espacios asociados a una matriz es:

   1) $R(A^t)$ = espacio renglón de A: dimensión = r
   2) $N(A)$ = espacio nulo de A: dimensión = $n-r$
   3) $R(A)$ = espacio columna de A: dimensión = r
   4) $N(A^t)$ = espacio nulo izquierdo de A: dimensión = $m-r$

   Tal que:

   $n = dim\ R(A) + dim\ N(A)$
   $m = dim\ R(A^t) + dim\ N(A^t)$

2. El sistema $Ax=b$ tiene al menos una solución x para toda b si y solo si las columnas de A generan R^m, entonces $r=m$. En este caso existe una inversa derecha H de $n\times m$ tal que $AH=I_m$, la matriz identidad de orden m. Esto es posible solo si $m=n$.

El sistema $Ax=b$ tiene a lo sumo una solución x para cada b si y solo si las columnas son linealmente independientes, entonces $r=n$. En este caso existe una inversa izquierda G de $n\times m$ tal que $GA=I_n$, la matriz identidad de orden n. Esto es posible solo si $m=n$.

3. Dos subespacios V y W del mismo espacio R^n son ortogonales si cada vector $v\in V$ es ortogonal a cada vector $w\in W$: $v^t w=0$ para toda v y w.

4. Dado un subespacio V de R^n, el espacio de todos los vectores ortogonales a V es el complemento ortogonal de V y se denota por $V^{\perp}$. Así se tiene entonces:

$$N(A) = \left(R(A^t)\right)^{\perp}, \qquad R(A^t) = N(A)^{\perp}$$
$$N(A^T) = \left(R(A)\right)^{\perp}, \qquad R(A) = \left(N(A^t)\right)^{\perp}.$$

5. Un producto interior en un espacio vectorial V es una función que a cada par de vectores x y y en V asocia un número real $<x, y>$de tal manera que se satisfacen los siguientes axiomas, para todos los vectores x, y, z en V y todos los escalares k.

1) $<x, y> = <y, x>$ axioma de la simetría
2) $<x+y, z> = <x, z> + <y, z>$ axioma de la aditividad
3) $<kx, y> = k <x, y>$ axioma de la homogeneidad
4) $<x, x> \geq 0$ $y$$<x, x> = 0 <==> x = 0$ axioma de positividad.

6. Es posible convertir cualquier conjunto de vectores independientes a_1, $a_2,\cdots,a_n$ en un conjunto de vectores ortogonales mediante el proceso de Gram-Schmidt: primero se fija $a_1=v_1$ después cada v_i es ortogonal a las $v_1,\cdots$, v_{i-1} precedentes:

$$v_1 = a_1 - \frac{< v_1, a_1 >}{\|v_1\|^2} v_1 - \cdots - \frac{< v_{i-1}, a_1 >}{\|v_{i-1}\|^2} v_2$$

7. Sea $Ax=b$ un sistema lineal de m ecuaciones con n incógnitas, donde $m>n$ (un sistema sobredeterminado) y el rango de A es n. La solución de mínimos cuadrados del sistema correspondiente $Ax\approx b$ de aproximaciones lineales es el vector $x=r$ que minimiza la magnitud del vector error.

8. Es decir, de todos los vectores Ar en el espacio columna el que minimiza $\|Ar-b\|$ es la proyección $bs=Ar$ de b en el espacio columna S. Entonces $Ar = A(A^tA)^{-1}A^t\, b$; y de aquí se tiene que el vector solución r que es óptimo está dado por: $r=(A^tA)^{-1}\, A^t\, b$

6.9. APLICACIONES DE LOS ESPACIOS VECTORIALES

En esta sección se presenta la aplicación a los mínimos cuadrados que es muy útil para interpolar datos, es decir para encontrar valores desconocidos con base en datos de una muestra. Se busca una función que mejor aproxime estos datos ya sea lineal o alguna curva.

MÉTODO DE LOS MÍNIMOS CUADRADOS

Supongamos que se obtienen datos de mediciones de la forma (a_i, b_i) mediante observación o experimentación y se ubican como puntos de datos en el plano x–y. Es conveniente hallar una relación matemática $Y=f(x)$ que represente razonablemente bien los datos, de manera que podamos efectuar predicciones de valores no medidos. Dependiendo de la naturaleza del experimento y de la configuración de los puntos de datos localizados, podemos decidir acerca de un tipo apropiado de función $y=f(x)$ como una función lineal, una función cuadrática o exponencial, entre otras. Se pueden presentar problemas como el siguiente:

De acuerdo con la ley de Hooke, la distancia que se estira un resorte es proporcional a la fuerza aplicada. Supongamos que colocamos 4 pesos distintos, a_1, a_2, a_3 y a_4 sucesivamente en la parte inferior de un resorte. Medimos las cuatro longitudes b_1, b_2, b_3 y b_4 del resorte estirado y suponga que se obtienen los datos de la siguiente tabla:

ai = peso en gramos	2.0	4.0	5.0	6.0
bi = longitud en cms.	2.5	4.5	7.0	8.5

Debido a la ley de Hooke, esperamos que los puntos de datos (a_i, b_i) estén cerca de alguna recta con ecuación:

$$y = f(x) = r_0 + r_1x$$

Donde r_0 es la longitud del resorte y r_1 es la constante del resorte. Esto es, si nuestras mediciones fueran exactas y el resorte ideal, tendríamos $b_i=r_0+r_1a_i$ para valores específicos r_0 y r_1.

Como solamente tenemos las dos incógnitas r_0 y r_1, bastarán dos mediciones para hallarlas, sin embargo, en la práctica esperamos tener algún error en las mediciones físicas, por lo tanto, hacemos más mediciones de las técnicamente necesarias con la esperanza de que en general, los errores se cancelen unos con otros. La sustitución de cada punto de datos (a_i, b_i) en la ecuación da una ecuación lineal con dos incógnitas r_0 y r_1, así los 4 puntos de datos del problema dan lugar a un sistema lineal de 4 ecuaciones con 2 incógnitas. Dicho sistema lineal con más ecuaciones que incógnitas se llama sobre determinado, y se espera que dicho sistema sea inconsistente, entonces debemos encontrar valores para las incógnitas r_0 y r_1 que estén los más cerca posible, de satisfacer las 4 ecuaciones. Geométricamente esto equivale a encontrar la recta en el plano que esté más cerca de pasar por los 4 puntos de datos.

En general, consideramos el problema de hallar una recta o función lineal $f(x) = r_0 + r_1x$ que ajuste "mejor" los datos (a_i, b_i) para $i=1, 2,\cdots, m$ donde $m>2$. Si no hubiera error en las mediciones y los datos fueran realmente lineales, entonces para algún r_0 y r_1 se tendría:

$$b_i = r_0 + r_1 a_i \qquad \text{para } i = 1,2,\cdots,m$$

Estas m ecuaciones en las 2 incógnitas r_0 y r_1 forman un sistema sobre determinado de ecuaciones que probablemente no tenga solución. Pero nuestros puntos satisfacen realmente un sistema de aproximaciones lineales que se pueden expresar como:

$$\begin{bmatrix} b_1 \\ b_2 \\ \gg \\ \cdots \\ b_n \end{bmatrix} \approx \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \cdots & \cdots \\ \cdots & \cdots \\ 1 & a_n \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix} \qquad (6.2)$$

O simplemente $b \approx Ar$.

Tratamos de encontrar un vector solución óptimo r para el sistema (6.2). Para cada vector r el vector error $Ar–b$ mide a qué distancia está el sistema (6.1) de un

sistema con solución r. Las normas de las componentes del vector $Ar-b$ representan las distancias $d_i = r_0 + r_1 a_i - b_i$ como se muestra en la figura 6.9

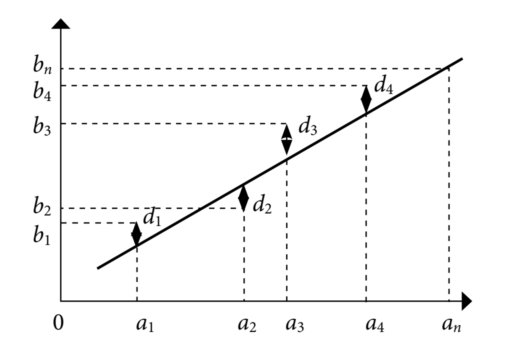


Figura 6.9 Distancias en el sistema

Queremos entonces minimizar nuestro vector de error $Ar-b$. Minimizaremos la longitud del vector error, pero esto equivale a minimizar $\|Ar-b\|^2$ o lo que es lo mismo, minimizar:

$$d_1^2 + d_2^2 + \cdots + d_m^2$$

De aquí el nombre de mínimos cuadrados.

Si a_1 y a_2 denotan las columnas de A en el sistema (6.2) entonces el vector $Ar = r_0 a_1 + r_1 a_2$ está en el espacio columna de A.

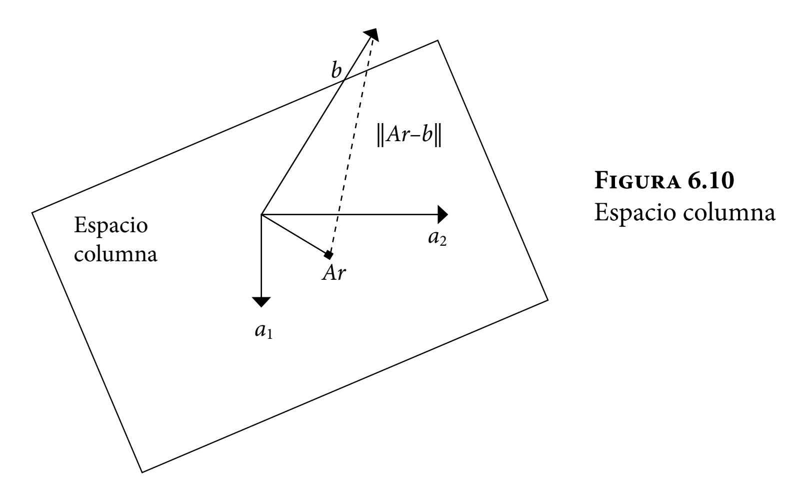


Figura 6.10 Espacio columna

Entonces de la figura 6.10, resulta claro que de todos los vectores Ar en el espacio columna el que minimiza $\| Ar-b \|$ es la proyección $b_s = Ar$ de b en el espacio

columna S. Entonces $Ar = A(A^tA)^{-1}A^tb$; y de aquí se tiene que el vector solución r que es óptimo está dado por:

$$r = (A^tA)^{-1} A^t b$$

Ahora sí, volviendo a los datos del problema, hacemos el ajuste de mínimos cuadrados por medio de una recta, esto es:

$$y = r_0 + r_1 x$$

Formamos el sistema $y \approx Ar$:

$$\begin{bmatrix} 2.5 \\ 4.5 \\ 7 \\ 8.5 \end{bmatrix} \approx \begin{bmatrix} 12 \\ 14 \\ 15 \\ 16 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^t A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 15 \\ 16 \end{bmatrix} = \begin{bmatrix} 4 & 17 \\ 17 & 81 \end{bmatrix}$$

$$A^t A = (1/35) \begin{bmatrix} 81 & -17 \\ -17 & 4 \end{bmatrix}$$

Entonces $r = (A^tA)^{-1} A^t b = [-0.9, 1.5]^t$.

De donde la recta que mejor ajusta los datos está dada por la ecuación $y = -0.9 + 1.5x$.

Ejemplo 6.26

Si se varía la carga que se aplica a una estructura y se mide la deformación que produce donde x es la carga, y es la lectura del medidor de la deformación. A menos que la carga sea tan grande que el material se haga plástico lo normal en la teoría de la elasticidad es la relación lineal $y = r_0 + r_1 x$. Dadas las siguientes mediciones encuentre la recta que mejor se ajusta.

x	0	1	3	4
y	0	1	2	5

Solución

Formamos el sistema $y=Ar$.

$$\begin{bmatrix} 0 \\ 1 \\ 2 \\ 5 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^t A = \begin{bmatrix} 4 & -17 \\ 17 & 81 \end{bmatrix}$$

$$(A^t A)^{-1} = (1/20) \begin{bmatrix} 13 & -4 \\ -4 & 2 \end{bmatrix}$$

Entonces $r = (A^tA)^{-1} A^t b = [-0.2, 1.1]^t$.

De donde la recta que mejor ajusta los datos está dada por la ecuación $y=-0.2 + 1.1x$.

Ejemplo 6.27

Una población de conejos de una gran isla se estimó todos los años desde 1981 hasta 1984 y se obtuvieron los datos que se listan en la siguiente tabla:

ai = año observado	1	2	3	4
bi = # de conejos en unidades de 1000	3	4.5	8	17
zi = Ln bi	1.1	1.5	2.08	2.83

Sabiendo que el crecimiento de la población es exponencial en ausencia de enfermedades, depredadores, hambre, etc. se espera que una función exponencial.

$$b_i = re^{sai} \quad \text{o} \quad b = re^{sa}$$

Sea la mejor representación de los datos. Encuentre la función exponencial que mejor ajusta los datos y con ella haga una proyección de la población de la isla para 1991.

Solución

Observe que usando logaritmos es posible convertir esta función exponencial a la forma lineal.

$$Ln\ b = Ln\ r + s(a)$$

De donde en la tabla se agrega el renglón de $z = ln\ (b_i)$.

Entonces como es lineal la función así escrita se usa

$$x = (A^tA)^{-1}A^t\ b$$

Donde:

$$\begin{bmatrix} 1.1 \\ 1.5 \\ 2.08 \\ 2.83 \end{bmatrix} \approx \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^t\ A = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}$$

$$(A^t\ A)^{-1} = \begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 1/5 \end{bmatrix}$$

Entonces $(A^tA)^{-1}\ A^t\ b = [0.435, 0.577]^t = [ln\ r, s]^t$.

Por lo tanto, $ln\ b = ln\ r + s(a_i) = 0.435 + 0.577a$ de donde:

$$b = re^{sa} = e^{0.435}\ e^{0.577a} = 1.54\ e^{0.577a}$$

En términos de funciones se tiene $f(x) = 1.54\ e^{0.577x}$.

A partir de esta función se puede proyectar la población de conejos para 1991 haciendo f(11) 1000 = 570778 conejos.

APROXIMACIÓN CUADRÁTICA

Ahora queremos ajustar una curva cuadrática a los n puntos de datos. Recuerde que una cuadrática en x es cualquier expresión de la forma:

$$y = a + bx + cx^2$$

Que representa a una parábola en el plano. Si los n puntos dados estuvieran en la parábola se tendría:

$$\begin{array}{l} y_1 = a + bx_1 + cx_1^2 \\ y_2 = a + bx_2 + cx_2^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = a + bx_n + cx_n^2 \end{array} \qquad (6.3)$$

Para:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad A = \begin{bmatrix} 1 & x_1 & x_{12} \\ \cdot\cdot & \cdot\cdot & \cdot\cdot \\ 1 & x_n & x_{n2} \end{bmatrix} \qquad r = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Entonces (6.3) se puede escribir como:

$$y = Ar$$

Igual que antes, si los puntos dados no están todos en la parábola entonces $y – Ar = 0$ y se tiene que para cualquier r nuestro problema vuelve a ser:

Encontrar un vector r en R^3 tal que $[y – Ar]$ sea mínimo y usando un razonamiento similar al anterior se tiene:

$$r = (A^t A)^{-1} A^t y$$

Ejemplo 6.28

En una reciente exhibición de yates se hicieron las observaciones listadas en la tabla siguiente, donde se relacionan los precios b_i de las embarcaciones y sus pesos a_i.

ai = peso en tons	2	4	5	8
bi = precio en unidades de 10 000	1	3	5	12

Al localizar los puntos de datos (a_i, b_i)como se muestra en la figura 6.11 podemos esperar que una función cuadrática de la forma:

$$y = f(x) = r_0 + r_1 x + r_2 x^2$$

Ajuste bien los datos:

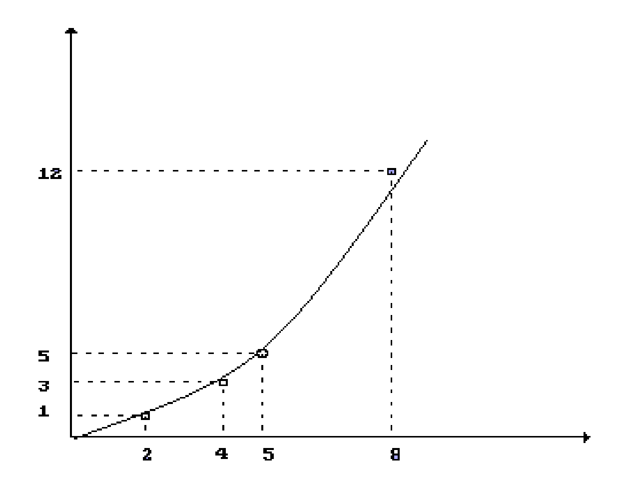


Figura 6.11 Ajuste de los datos por una función cuadrática

Solución

Formamos el sistema $y=Ar$.

$$\begin{bmatrix} 1 \\ 3 \\ 5 \\ 12 \end{bmatrix} \approx \begin{bmatrix} 1 & 2 & 14 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 8 & 64 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$A^t A = \begin{bmatrix} 4 & 19 & 109 \\ 19 & 109 & 709 \\ 109 & 709 & 4993 \end{bmatrix}$$

$$(A^t A)^{-1} = (1/5400) \begin{bmatrix} 12744 & -3348 & -6624 & 2628 \\ -4464 & 2538 & 3744 & -1818 \\ 360 & -270 & -360 & 270 \end{bmatrix}$$

Entonces $y = (A^t A)^{-1} A^t b = [0.207, 0.01, 0.183]^t$.

Así la función cuadrática que mejor aproxima los datos en el sentido de los mínimos cuadrados es:

$$y = 0.207 + 0.01x + 0.183x^2$$

Para graficar esta función se hace:

a_i	b_i	$f(a_i)$
2	1	0.959
4	3	3.175
5	5	4.832
8	12	11.999

Como quedó en la figura 6.11

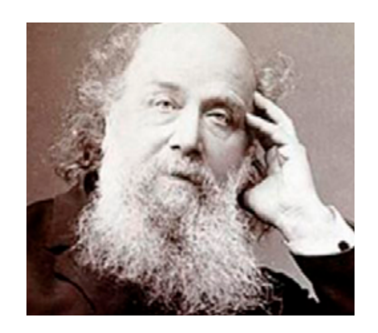
6.10. NOTAS HISTÓRICAS



EL RANGO DE UNA MATRIZ lo definió en 1879 **Georg Frobenius** (1849-1917) como sigue: si se anulan todos los determinantes de grado $(r+1)$, pero no todos los de grado r, entonces r es el rango de la matriz. Frobenius usó este concepto para tratar las cuestiones de formas canónicas para ciertas matrices de enteros y las soluciones de ciertos sistemas de congruencias lineales.

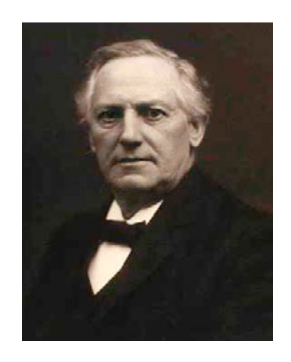
Por otro lado, **James Sylvester** definió en 1884 la nulidad para matrices cuadradas, como sigue: la nulidad de una matriz de $n \times n$ es i si todo menor (determinante) de orden $n-i+1$ (y, por tanto, de todo orden superior) es igual a 0 e i es el mayor de los números para los cuales esto es cierto.

Aquí, Sylvester estaba interesado, como en buena parte de su carrera matemática, en descubrir invariantes, esto es, propiedades de objetos matemáticos particulares que no cambian bajo tipos específicos de transformación. Él procedió a probar lo que llamó una de las leyes cardinales en la teoría de matrices: que la nulidad del producto de dos matrices no es menor que la nulidad de cualquier factor ni mayor que la suma de las nulidades de los factores.

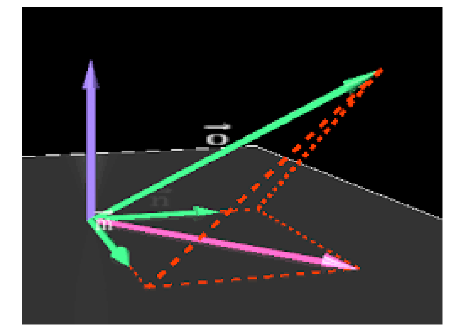


EL PROCESO DE GRAM-SCHMIDT debe su nombre al matemático danés **Jørgen P. Gram** (1850-1916), y al alemán **Erhard Schmidt** (1876-1959). Lo publicó primero Gram en 1883 en un artículo titulado "Desarrollo de series usando el método de los mínimos cuadrados", y Schmidt lo publicó de nuevo en 1907, con una cuidadosa demostración, en un trabajo sobre ecuaciones integrales. De hecho, Schmidt incluso se refirió al trabajo de Gram. Para Schmidt igual que para Gram, los vectores eran funciones continuas definidas en un intervalo [a,b] con el producto interno de dos funciones *s* y *F* dado por la integral de su producto en ese intervalo. Sin embargo, Schmidt fue más explícito que Gram al escribir el proceso con gran detalle y probar que el conjunto de funciones F_i derivado de su conjunto original s_i, era de hecho, un conjunto ortonormal.

Schmidt, que estuvo en la Universidad de Berlín desde 1917 hasta su muerte, es más conocido por su trabajo decisivo sobre los espacios de Hilbert, espacios de sucesiones de números complejos de cuadrado sumable. De hecho, aplicó el proceso de Gram-Schmidt a conjuntos de vectores en esos espacios para ayudar a desarrollar condiciones necesarias y suficientes para que dichos conjuntos sean linealmente independientes.







LA PSEUDOINVERSA DE MOORE PENROSE

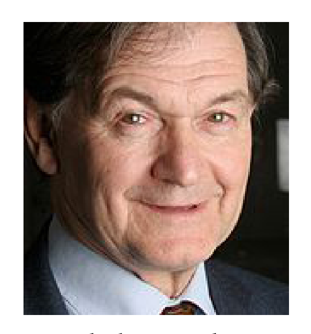
En álgebra lineal, una matriz A^+ pseudoinversa de una matriz A es una generalización de la matriz inversa.

El tipo de matriz pseudoinversa más conocido es la inversa Moore–Penrose, que fue descrita independientemente por **Eliakim Hastings Moore** en 1920, Arne Bjerhammar en 1951, y **Roger Penrose** quien, en 1955 siendo todavía un

estudiante reinventó la inversa generalizada. Anteriormente, Erik Ivar Fredholm introdujo el concepto de pseudoinversa de operadores integrales en 1903. Cuando se refiere a una matriz, el término pseudoinversa, sin más especificación, se usa a menudo para indicar la inversa Moore-Penrose. El término inversa generalizada se usa a veces como un sinónimo de pseudoinversa.

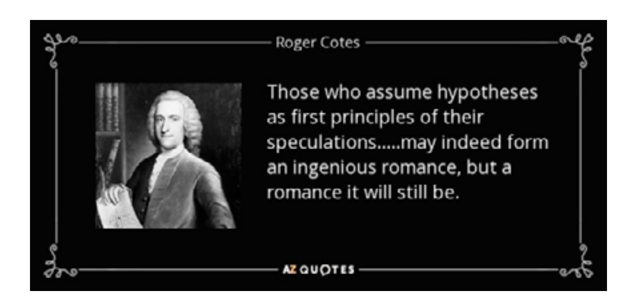


UNA TÉCNICA MUY CERCANA A LA DE LOS MÍNIMOS CUADRADOS es la que desarrolló Roger Cotes (1682-1716), el genial matemático que editó la segunda edición de los **Principia de Isaac Newton**, en una obra que trataba errores en las observaciones astronómicas, escrita alrededor de 1715.



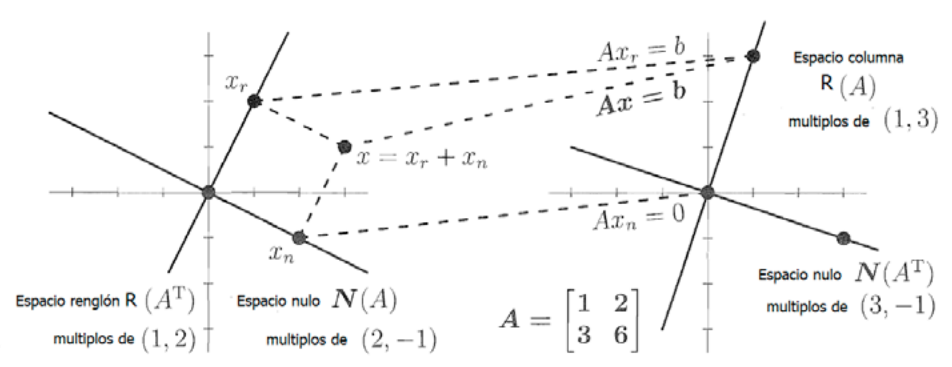
Sin embargo, fue Carl Gauss, a los 16 años el primero en formular el principio completo, mientras ajustaba aproximaciones relacionadas con la distribución de los números primos. Más tarde Gauss, afirmo que durante años había usado a menudo el método, por ejemplo, en sus cálculos sobre las órbitas de asteroides. Gauss publicó el método en 1809 e hizo una exposición definitiva 14 años después.

Por otro lado, debemos a Adrien-Marie Legendre (1752-1833), fundador de la teoría de las funciones elípticas, la primera publicación del método de los mínimos cuadrados en un trabajo de 1806, sobre la determinación de las órbitas de cometas. Después de la publicación de Gauss en 1809, Legendre le escribió censurándolo por reclamar el método como propio. Todavía en 1827, Legendre seguía acusando a Gauss de apropiarse de los descubrimientos de otros. De hecho, el problema radicaba en el fallo de Gauss de no publicar a tiempo sus descubrimientos; solamente los mencionaba cuando ya habían sido publicados por otros.



6.11. EJERCICIOS PROPUESTOS

1. Hemos visto que los espacios fila $R(A^t)$ y columna $R(A)$ tienen la misma dimensión r, y que los espacios $R(A^t)$ y $N(A)$ son complementos ortogonales así como también los espacios $R(A)$ y $N(A^t)$. Un vector x arbitrario se puede descomponer como $x = x_r + x_n$ y la matriz A transforma a la componente x_r del espacio fila en un vector $Ax_r = Ax$ en el espacio columna, mientras que transforma a la componente x_n del espacio nulo en cero, como se ilustra en la siguiente figura:



a) Para la siguiente matriz A, dado $x = (3, 3, 3)$ descomponga en $x = x_r + x_n$. (Sugerencia: considere las bases de los subespacios renglón y nulo para expresar a x como una combinación lineal de ellos).

$$A=\begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix}$$

b) Verifique que $Ax=Ax_r$ en el inciso *a*.

2. a) Encuentre el complemento ortogonal del plano generado por los vectores (1,1,2) y (1,2,3) considerándolos como las filas de A y resolviendo $Ax=0$. Recuerde que el complemento es toda una recta.

   b) Construya una ecuación homogénea en tres incógnitas cuyas soluciones son las combinaciones lineales de los vectores (1,1,2) y (1,2,3).

3. **Matrices de Incidencia.** Desde hace mucho tiempo, el análisis de redes ha jugado un papel muy importante en la ingeniería eléctrica. Sin embargo, se ha visto en las últimas décadas que ciertos conceptos y herramientas de la

teoría de redes son útiles también en otros contextos, por ejemplo, sistemas de comunicación, programas de producción etc.

Usando la terminología de teoría de gráficas, una gráfica consiste en un conjunto de puntos de unión llamados nodos, unidos por rectas llamadas arcos. Se considera que una red es una gráfica con un flujo de algún tipo en los arcos. La matriz de incidencia es la siguiente:

$$A=\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Los arcos corresponden a las columnas de la matriz y los nodos a los renglones, el primer arco va del nodo 1 al nodo 2 y en la columna 1 se tiene 1 en el primer renglón y −1 en el segundo renglón, las otras columnas se construyen de la misma manera. En cada renglón se indica el nodo que se deja por 1 y el nodo que entra por −1. Entonces A es la matriz nodos-arcos de la gráfica.

Dibuje la red asociada a esta matriz, encuentre una base para cada uno de los cuatro espacios asociados a la matriz A y verifique si hay una representación en la red original para dichas bases.

4. **Leyes de Kirchhoff.** Una aplicación inmediata de redes se da en el caso en que el flujo que corre por sus arcos es corriente eléctrica. Considere la red de la figura siguiente, si fluye corriente directa por la red entonces la única manera de mantener el equilibrio es satisfacer las dos leyes de Kirchhoff:

   Primera Ley: En cada nodo, la suma de las corrientes de entrada es igual a la suma de las corrientes de salida. Por ejemplo: $I_3 = I_1 + I_4$ en el nodo 1 de la figura anterior.

   Segunda Ley: La suma de caídas de voltaje es cero en cada circuito cerrado. Si la caída es E_k en la dirección indicada por I_k, entonces la segunda ley requiere que $E_4 + E_6 + E_3 = 0$ alrededor del circuito con nodos 1–4–3.

   Estas leyes en términos del álgebra matricial solo dependen de la manera en que los nodos están ligados por los arcos, y de la dirección de las flechas, pero no del tamaño de las resistencias en la red. Las relaciones estructurales entre los nodos están completamente descritas por la matriz de incidencia:

$$M=\begin{bmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 & -1 \end{bmatrix} \begin{matrix} \textit{nodo } 1 \\ \textit{nodo } 2 \\ \textit{nodo } 3 \\ \textit{nodo } 4 \end{matrix}$$

En la primera fila de M hay un -1 para indicar la entrada del arco 3 y hay $+1$ para indicar la salida de los arcos 1 y 4. Recordemos que la primera ley de Kirchhoff da $I_3 = I_1 + I_4$, esto se traslada inmediatamente a la forma matricial: si I es el vector columna formado por las seis corrientes, entonces $MI=0$. Este es un sistema de cuatro ecuaciones con seis incógnitas.

Para interpretar la segunda ley, fijaremos el potencial del nodo 1 en $p_1=0$ y definimos p_i en los demás nodos de tal manera que la caída de voltaje del nodo i al nodo j (si hay un arco en esa dirección), sea igual a la diferencia de los potenciales $E=p_i-p_j$. La segunda ley garantiza que la caída total alrededor de un circuito es cero de modo que volvemos al mismo potencial con el que comenzamos. En términos de M, la caída p_i-p_j a través de cualquier arco viene de multiplicar el vector potencial p por la columna correspondiente a ese arco, por construcción esa columna tiene $+1$ en el nodo i y -1 en el nodo j. En otras palabras, las caídas de voltaje E son las componentes de M^Tp.

Enunciando brevemente estas leyes tenemos que I está en el espacio nulo de M y E está en su espacio fila. Como estos dos espacios son ortogonales para cualquier matriz M, tenemos el Teorema de Tellegen en teoría de circuitos: $E^TI=0$.

a) Encuentre las caídas de voltaje a través de cada arco, así como los potenciales p_2, p_3,p_4 dado que $p_1=0$. Verificar que $E^TI=0$.
b) Dibuje la red cuya matriz de incidencia es:

$$M=\begin{bmatrix} 1 & 0 & 0 & -1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 & -1 \end{bmatrix}$$

Con una batería de 6 volts en el arco 1–3 y resistencias unitarias en todos los arcos ¿Cuáles son las corrientes I y las caídas de voltaje E?

5. Sea X un espacio vectorial. Se dice que X tiene un producto interior, si asigna un número real $<x,y>$ a cada par de elementos x,y de X y se satisfacen los cuatro axiomas que ya se han visto en clase. Demuestre que en el plano o espacio vectorial R^2 las siguientes funciones son productos interiores.

   a) $<x,y> = 9x_1 y_1 + x_2 y_2$
   b) $<x,y> = x_1 y_1 + x_2 y_1 + x_1 y_2 + 2x_2 y_2$
   Donde x, y son elementos de R^2.

6. a) Considere a R^2 con el producto interior $<x,y> = 3x_1 y_1 + 2x_2 y_2$ donde $x=(x_1, x_2)$, $y=(y_1, y_2)$. Encuentre $\| z \|$ cuando:
   i. $z = (-1,3)$
   ii. $z = (6, 7)$

   b) Ahora usando el mismo producto interior del inciso a. determine $d(x,y)$ cuando:
   i. $x = (-1,2)$, $y = (2,5)$
   ii. $x = (3,9)$, $y = (3,9)$

7. Encuentre la proyección ortogonal de a sobre b si:

   a) $a = (2,1)$, $b = (-3,2)$
   b) $a = (-7,1,3)$, $b = (5,0,1)$

8. a) Demuestre que para cualesquiera dos vectores a y b de R^n se cumple $\|a + b\| \leq \|a\| + \|b\|$.
   b) Usando la desigualdad del inciso anterior demuestre también que $\|a - b\| \leq \|a\| + \|b\|$.

9. Propiedades de una matriz de proyección P. La matriz de proyección P_s de un subespacio S de R^n es idempotente y simétrica. Recíprocamente, toda matriz de orden $n \times n$ que sea idempotente y simétrica es una matriz de proyección, a saber, es la matriz de proyección para su espacio columna. Estas propiedades se pueden escribir como sigue:

   i. $P^2 = P$ (*P es idempotente*)
   ii. $P^T = P$ (*P es simétrica*)

Demuestre que la matriz de proyección $P = A(A^T A)^{-1} A^T$ satisface las dos condiciones anteriores.

10. a) Halle el vector proyección de (1,2,1) sobre el subespacio S de R^3 generado por los vectores $\{[3,1,2],[1,0,1]\}$.
    b) Halle el vector proyección de (1,0,0) en el subespacio S de R^3 generado por los vectores $\{[2,1,1],[1,0,2]\}$.

11. Considere a R^3 con el producto interior euclideano. Utilice el proceso de Gram-Schmidt para transformar la base $\{u_1, u_2, u_3\}$ en una base ortonormal.

    a) $u_1 = (1,1,1)$, $u_2 = (-1,1,0)$, $u_3 = (1,2,1)$
    b) $u_1 = (1,0,0)$, $u_2 = (3,7,-2)$, $u_3 = (0,4,1)$

Capítulo 7
VALORES Y VECTORES PROPIOS

7.1. INTRODUCCIÓN

En este capítulo veremos lo que se puede considerar la segunda parte de teoría de matrices, para ello es necesario entender lo que son los valores propios y cómo pueden ser útiles. Una de sus aplicaciones es la solución de un sistema de ecuaciones diferenciales ordinarias, donde al parecer la clave se encuentra en sus valores y vectores propios, también se pueden llamar característicos. Quizá el ejemplo más sencillo es el de los soldados que cruzan un puente. Cada vez que lo hacen dejan de marchar y lo cruzan solo caminando. La razón es que podrían marchar a una frecuencia igual a uno de los valores propios del puente y entonces comenzaría a oscilar. Un ingeniero trata de que las frecuencias naturales de su puente o de su cohete estén alejadas de las del viento o el chapoteo de la gasolina. En el otro extremo, un corredor de bolsa pasa su vida tratando de estar en correspondencia con las frecuencias naturales del mercado. Los valores y los vectores propios son los rasgos más importantes de prácticamente cualquier sistema dinámico.

Otro ejemplo interesante es la rotación de la Tierra. Siempre habrá alguna dirección que permanezca sin cambio, a saber, el eje de rotación. No necesariamente es aquel alrededor del cual la Tierra gira en la realidad, pero deberá haber algún polo norte y polo sur que permanezcan fijos. Estos polos son los vectores propios, con valores propios iguales a 1. En general todos los otros puntos se mueven y no vemos más vectores propios. Las únicas excepciones suceden cuando la rotación es de 360° (con vectores propios por donde sea) o de 180°. En el caso de 180°, el plano del ecuador está lleno de vectores propios; cada dirección del plano se ha invertido y el valor propio es –1. Este plano ecuatorial es un caso de un “espacio propio bidimensional”; el único valor propio $l=1$ tiene dos vectores propios independientes, por lo tanto, un plano de vectores propios. En general los valores propios no son ±1 y los vectores usualmente se estiran o

comprimen. Pero lo más importante es que tomando una rotación en el espacio tridimensional R^3, y exceptuando 180° y 360°; solo hubo una recta de vectores propios cuando esperábamos tres. Como se puede ver la gama de aplicaciones del tema que aborda este capítulo es muy amplia y es conveniente hacer algunas definiciones.

En este capítulo se usará indistintamente valores y vectores propios o característicos.

7.2. VALORES Y VECTORES PROPIOS

Sea $T: V \rightarrow V$ una transformación lineal. En una gran variedad de aplicaciones resulta útil encontrar un vector $v \in V$ tal que Tv y v sean paralelos. Esto es, buscamos un vector v y un escalar a tal que:

$$Tv = \alpha v \tag{7.1}$$

Si $v \neq 0$ y α satisface (7.1), entonces α se conoce como un valor propio o característico de T y v es un vector propio o característico de T correspondiente al valor propio α.

> ***Definición 7.1***
> Si A es una matriz de $n \times n$, entonces un escalar α es un valor propio de A sí hay un vector v distinto de cero tal que:
>
> $$Av = \alpha v$$
>
> El vector $v \neq 0$ se llama un vector propio de A correspondiente a un valor propio α.

Ejemplo 7.1

Sea A tal que:

$$A = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}$$

Entonces:

$$A=\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}10 & -18\\ 6 & -11\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}$$

De esta forma $\alpha_1=1$ es un valor propio (característico) de A, correspondiente al vector propio (característico) $v_1=[2, 1]^t$, análogamente:

$$A=\begin{bmatrix}3\\2\end{bmatrix}=\begin{bmatrix}10 & -18\\ 6 & -11\end{bmatrix}\begin{bmatrix}3\\2\end{bmatrix}=\begin{bmatrix}-6\\-4\end{bmatrix}=-2\begin{bmatrix}3\\2\end{bmatrix}$$

Por lo que $\alpha_2= -2$ es un valor propio de A con su correspondiente vector propio $v_2=[3, 2].^t$

Sea α un valor propio de A, entonces existe un vector no nulo $v=(x_1, \cdots, x_n)^t \neq 0$, tal que,

$$\begin{aligned}&Av = \alpha v = \alpha I v\\ &(A - \alpha I)v = 0 \qquad (7.2)\end{aligned}$$

Si A es de orden, entonces la ecuación (7.2) es un sistema homogéneo de n ecuaciones con n incógnitas, como por hipótesis (ya que v es un vector propio distinto de cero), el sistema tiene soluciones no triviales, entonces: $\det(A-\alpha I)=0$

A esta ecuación se le denomina la **ecuación característica de A.**

Si $A=(a_{ij})$ entonces la ecuación anterior se puede escribir

$$\begin{bmatrix}a_{11}-\alpha & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-\alpha & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & \cdots & \cdots & a_{nn}-\alpha\end{bmatrix}=0$$

Si desarrollamos el determinante obtenemos una expresión polinomial $p(\alpha)$ de grado n con coeficientes que incluyen a_{ij}, es decir:

$$\det(A-\alpha I) = p(\alpha)$$

El polinomio $p(\alpha)$ es el polinomio característico de la matriz A. Los valores propios de A son precisamente las soluciones de la ecuación característica $p(\alpha) = 0$.

Ejemplo 7.2

Sea $A=I$, entonces para todo vector $v \in R^n$, $Av=Iv=v$. De esta manera I es el único valor característico de A y cada vector v es un vector característico de I.

Ejemplo 7.3

Encuentre los vectores propios de la matriz A:

$$A=\begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$$

Solución

El polinomio característico de A es:

$$det\,(A - \alpha I) = \begin{vmatrix} 3 & -\alpha & 2 \\ 2 & & -\alpha \end{vmatrix} = \alpha^2 - 3\alpha - 4 = 0$$

$a^2 - 3\alpha - 4 = (\alpha-4)(a+1) = 0$, de donde, $\alpha_1 = -1$ y $\alpha_2 = 4$ son valores propios de A.

Ejemplo 7.4

Encuentre los valores propios de la matriz A:

$$A=\begin{bmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 3 & 1 \end{bmatrix}$$

Solución:

$$\det\,(A-\alpha I) = \begin{vmatrix} 2-\alpha & 1 & 0 \\ -1 & -\alpha & 1 \\ 1 & 3 & 1-\alpha \end{vmatrix} = -(\alpha-2)\,(\alpha^2-\alpha-2) = -\,(\alpha-2)\,(\alpha-2)\,(\alpha+1)$$

$-(\alpha-2)\,(\alpha-2)\,(\alpha+1) = 0$ de donde, $\alpha_1=-1$ y $\alpha_2= \alpha_3= 2$ son valores propios de A.

En general si se tiene una matriz A de orden $n \times n$

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = 0$$

Entonces:

$$p(\alpha) = \det(A - \alpha I) = \begin{vmatrix} a_{11}-\alpha & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\alpha & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn}-\alpha \end{vmatrix}$$

Y $p(\alpha)$ se puede escribir como:

$$p(\alpha) = \alpha^n + b_{n-1}\alpha^{n-1} + \cdots + b_1\alpha + b_0 = 0 \qquad (7.3)$$

La ecuación (7.3) tiene n raíces, varias de las cuales pueden repetirse. Si $a_1, \cdots, a_m$ son las diferentes raíces de (7.3) con multiplicidades $r_1, r_2, \cdots, r_m$, entonces (7.3) se puede factorizar para obtener:

$$p(\alpha) = (\alpha - \alpha_1)^{r1} (\alpha - \alpha_2)^{r2} \cdots (\alpha - \alpha_m)^{rm} = 0$$

Los números $r_1, r_2, \cdots, r_m$ se llaman **multiplicidades algebraicas** de los valores propios de $a_1, \cdots, a_m$.

Ejemplo 7.5

Todo polinomio de grado n con coeficientes reales o complejos tiene n raíces exactamente (contando multiplicidades), por ejemplo; el polinomio $(\alpha - I)^5$ tiene cinco raíces todas iguales al número uno. Puesto que todo valor característico de A es una raíz de la ecuación característica de A se tiene:

> ***Definición 7.2***
> Si contamos multiplicidades, cada matriz de tiene exactamente n valores característicos.

Teorema 7.1

Si A es una matriz de orden $n \times n$, entonces las siguientes afirmaciones son equivalentes:

a) α es un valor característico de A.
b) El sistema $(A-\alpha I)v = 0$ tiene soluciones no triviales.
c) Existe un vector $v \in R^n v \neq 0$, tal que $Av=av$

Si α es un valor característico de A, entonces el espacio solución del sistema de ecuaciones $(A -\alpha I)v=0$ se denomina el espacio característico de A correspondiente a α, y los vectores diferentes de cero en el espacio característico, se denominan los vectores característicos de A correspondientes a α.

INTERPRETACIÓN GEOMÉTRICA

Si v es un vector característico de A correspondiente aa,entonces $Av = av$. Por lo tanto, la multiplicación por A mapea a v en un múltiplo escalar de sí mismo; por consiguiente, dependiendo del valor de$a$, la multiplicación por A dilata a v, lo contrae o invierte su dirección, ver figura 7.1.

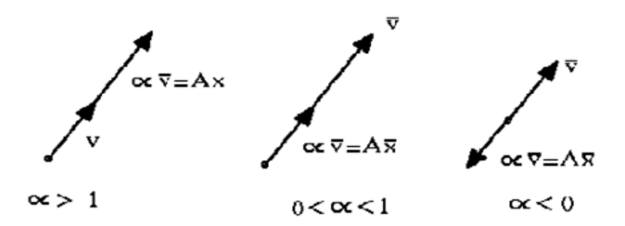


FIGURA 7.1 El vector v afectado por los valores de α

Ejemplo 7.6

Sea la matriz A como sigue:

$$A=\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$$

Encuentre los valores y los vectores propios.

Solución

Sea:

$$det\ (A-\alpha) = \begin{vmatrix} 4-\alpha & 1 \\ 0 & 4-\alpha \end{vmatrix} = (\alpha-4)^2 = 0$$

α=4 es el valor característico de multiplicidad 2, de donde:

$$(A-4I)\ V = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ 0 \end{bmatrix}$$

Lo que implica que $v_1=[1, 0]^t$ es un vector propio.

Ejemplo 7.7

Sea la matriz A como sigue:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

Encuentre los valores y vectores característicos de A.

Solución

Sea

$$det\ (A-\alpha I) = \begin{vmatrix} 4-\alpha & 0 \\ 0 & 4-\alpha \end{vmatrix} = (\alpha-4)^2 = 0$$

$\alpha=4$ es el valor característico de multiplicidad 2, de donde: $Av=4v$. Lo que implica que $v_1=[1, 0]$ y $v_2 = [0, 1]$ generan el espacio característico de A.

> ***Definición 7.3***
> Sea a un valor característico de A. El subespacio E_a, se denomina el espacio característico de A correspondiente al valor característico a.

Teorema 7.2

Sea α un valor característico de la matriz A de orden $n\times n$ y sea $E_a=\{v : Av = \alpha v\}$, entonces E_a es un subespacio de R^n (en general de E^n).

Demostración

Si $Av=av$, entonces $(A-\alpha I)\ v=0$; de donde, E_a es el espacio nulo de la matriz $A-\alpha I$, la cual es un subespacio de R^n.

Teorema 7.3

Sea A una matriz de $n\times n$ y sean $\alpha_1, \alpha_2, \cdots, \alpha_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $v_1, v_2, \cdots, v_m$. Entonces v_i son linealmente independientes o bien, los vectores característicos correspondientes

a valores característicos diferentes son linealmente independientes. (La demostración se hace por inducción).

Ejemplo 7.8
Sea la matriz A, como sigue:

$$A=\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Encuentre los valores y vectores propios, así como el espacio característico correspondiente.

Solución
Sea $\det(A-\alpha I) = -\alpha^3+ 6\alpha^2+ 15\alpha+ 8 = -(\alpha+1)^2 (\alpha-8) = 0$

Lo que implica $\alpha_1 = 8$ y $\alpha_2 = -1$con multiplicidad algebraica 2:

Para $\alpha_1= 8$ se obtiene:

$$v_1= [2, 1, 2]^t \text{ y } E_8 = \{[2, 1, 2]\}.$$

Para $\alpha_2= -1$ se obtiene:

$$v_2 = [1, -2, 0]^t \text{ y } v_3 = [0, -2, 1]^t$$

Con

$$E_{-1}= \left\{\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}\begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}\right\}$$

En los ejemplos que hemos visto, encontramos un valor característico con una multiplicidad algebraica de dos o más. Sin embargo, el número de vectores característicos linealmente independientes no necesariamente es igual a la multiplicidad algebraica del valor característico (como en el ejemplo 7.6). Esto se puede corroborar con la siguiente definición.

Definición 7.4

Sea α un valor característico de A, entonces la multiplicidad geométrica de a es la dimensión del espacio característico correspondiente a α (que es la nulidad de la matriz $A - \alpha I$) esto es:

Multiplicidad geométrica de α = dim E_α

Si A es una matriz de 2×2 y α un valor característico con multiplicidad algebraica 2, entonces la multiplicidad geométrica de α es menor o igual a 2; puesto que puede haber al menos dos vectores linealmente independientes en un espacio de dos dimensiones.

Si A es de orden 3 con dos valores característicos α_1 y α_2 de multiplicidad algebraica 1 y 2 respectivamente, entonces la multiplicidad geométrica de α_2 es menor o igual a 2.

Teorema 7.4

Sea α un valor característico de A, entonces:

Multiplicidad geométrica de $\alpha \leq$ multiplicidad algebraica de α

7.3. DIAGONALIZACIÓN

Teorema 7.5

Sea A una matriz de orden $n \times n$, entonces A tiene n vectores característicos linealmente independientes si y solo si la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, A tiene n vectores característicos linealmente independientes si todos sus valores característicos son distintos (pues si no, podrían ser menores que n).

Como ya se ha visto en varios ejemplos, algebraicamente y desde el punto de vista de los cálculos, el problema de los valores propios es mucho más difícil que $Ax=b$; sin embargo, se pueden usar algunos resultados para que sea más fácil trabajar con ellos.

Definición 7.5

La suma de los n valores propios, es igual a la suma de las n entradas de la diagonal de A:

$a_1+\cdots+a_n= a_{11} +\cdots+a_{nn}$

Esta suma es la **traza** de A. Además, el producto de los n valores propios es igual al determinante de A.

Ejemplo 7.9

Sea la matriz simétrica:

$$A=\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Con valores propios $\alpha_1=0$, $\alpha_2=1$ y $\alpha_3=3$ entonces:

$$0 + 1 + 3 = 1 + 2 + 1 = 4 \text{ y } \det A = 0$$

Nota: No deben confundirse los valores propios de una matriz y sus entradas diagonales. Normalmente son completamente diferentes; sin embargo, se tiene la siguiente definición:

Definición 7.6

Si la matriz A es triangular (puede ser superior y en particular diagonal), entonces los valores propios $a_1,\cdots,a_n$ son exactamente los mismos que las entradas de la diagonal $\alpha_{11},\cdots,\alpha_{nn}$.

Ejemplo 7.10

Sea la matriz A como sigue:

$$A=\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Su polinomio característico es:

$$det\begin{vmatrix} 1-\alpha & ¼ & 0 \\ 0 & ¾-\alpha & 0 \\ 0 & 0 & ½-\alpha \end{vmatrix} = (1-\alpha)\,(¾-\alpha)\,(½-\alpha)$$

El determinante es el producto de las entradas diagonales. Obviamente, las raíces son:

$a_1 = 1$, $a_2 = ¾$ y $a_3 = ½$;
los valores propios ya estaban colocados a lo largo de la diagonal principal.

Este ejemplo, en el cual, se pueden encontrar los valores propios por inspección, señala lo más importante de este tema: **transformar a una matriz *A* en una matriz diagonal o triangular sin cambiar sus valores propios.** (La factorización *LU* no sirve para este caso, ya que los valores propios de *U* no son los valores propios de *A*).

Otra situación en que los cálculos son fáciles es: si ya tenemos los valores y vectores propios de una matriz *A*, entonces los valores propios de A^2 son exactamente $a_1^2, \cdots, a_n^2$ y cada vector propio de *A* es también un vector propio de A^2.

Teorema 7.6

Supongamos que la matriz *A* de tiene *n* vectores propios linealmente independientes; entonces, si se eligen estos vectores como las columnas de una matriz *S*, se sigue que: $S^{-1}AS$ es una matriz diagonal *A* con los valores propios de *A* en su diagonal.

$$S^{-1}AS = \Lambda = \begin{bmatrix} \alpha_1 & & \\ & \alpha_2 & \cdots \\ & \vdots & \alpha_n \end{bmatrix}$$

Demostración

Colocamos los vectores propios x_i en las columnas de *S* y se calcula el producto *AS*, una columna a la vez:

$$AS = A\begin{bmatrix} \vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots \\ x_1 & x_2 & \cdots & x_n & = & a_1x_1 & a_2x_2 & a_3x_n \\ \vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots \end{bmatrix}$$

Entonces:

$$\begin{bmatrix} \vdots & \vdots & \vdots \\ \alpha_1 x_1 & \alpha_2 x_2 & \alpha_3 x_n \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ x_1 & x_2 \cdots & x_n \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \alpha_1 & & \\ & \alpha_2 & \cdots \\ & \vdots & \alpha_n \end{bmatrix} = S\Lambda$$

$$AS = S\Lambda \qquad S^{-1}AS = \Lambda A = S\Lambda S^{-1}$$

La matriz S es invertible, ya que supusimos que sus columnas (los vectores propios son linealmente independientes).

Observación 1

Si la matriz A no tiene valores propios repetidos (los números $\alpha_1, \cdots, \alpha_n$ son distintos), entonces los n vectores propios son independientes y; por lo tanto, cualquier matriz con valores propios distintos puede diagonalizarse.

Observación 2

La matriz diagonalizada S no es única, ya que un vector propio x puede multiplicarse por una constante y seguir siendo vector propio; por lo tanto, se pueden multiplicar las columnas de S por cualquier constante distinta de cero y producir una nueva diagonalización S.

Observación 3

La ecuación $AS=S\Lambda$ es válida si las columnas de S son los vectores propios de A y no de otra manera. Supongamos que la primera columna de S es algún vector y, entonces la primera columna de SA es $\alpha_1 y$. Si esto va a corresponder con la primera columna de AS, que debido a la multiplicación de matrices es Ay, entonces y debe ser un vector propio $Ay= \alpha_1 y$.

Observación 4

No todas las matrices poseen n vectores propios linealmente independientes y por lo tanto, no todas las matrices son diagonalizables.

Ejemplo 7.11

Sea la matriz A como sigue:

$$A= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Sus valores propios son: $a_1 = a_2 = 0$, ya que es triangular

$$det\,(A-\alpha I) = \begin{vmatrix} -\alpha & 1 \\ 0 & -\alpha \end{vmatrix} = \alpha^2$$

Si x es un vector propio entonces:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ ó } x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

Aunque $\alpha=0$ es un valor propio de multiplicidad algebraica 2, solo tiene un espacio unidimensional de vectores propios. La multiplicidad geométrica de este valor propio es 1 y no podemos construir S.

Como $a_1=a_2=0$, A tendría que ser la matriz de ceros; pero si $S^{-1}AS=0$, entonces: $SS^{-1}AS=0$; lo que implica que $ASS^{-1}=0$, de aquí se deduce que $A=0$; lo que resulta imposible. Por lo tanto, no existe S tal que $S^{-1}AS=A$

Algunas matrices con valores propios repetidos pueden diagonalizarse. Si se repite un valore propio α, m veces, entonces para que se pueda diagonalizar A debe haber m vectores propios correspondientes.

Ejemplo7.12

Diagonalice la siguiente matriz A:

$$A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Solución

Los valores propios de A son: $\alpha_1=1$ y $\alpha_2=5$.

Para $\alpha_2=5$ se tienen los vectores propios:

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Y para $\alpha_1=1$ se tiene:

$$v_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

De donde; $\{v_1, v_2, v_3\}$ es un conjunto de vectores linealmente independientes entonces:

$$S = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Diagonaliza a A.

$$S = \begin{bmatrix} -½ & ½ & 0 \\ 0 & 0 & 1 \\ ½ & ½ & 0 \end{bmatrix} \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

El orden de las columnas de S no altera los valores propios solo su orden.

Teorema 7.7

Una matriz A de $n \times n$ es diagonalizable, si y solo si, la multiplicidad algebraica de cada valor propio es igual a su multiplicidad geométrica.

7.4. PROPIEDADES DE VALORES Y VECTORES PROPIOS

Sea la matriz A de orden $n \times n$

1. Si a es un valor propio de A con v como vector propio correspondiente, entonces Λ^k es un valor propio de A^k, de nuevo con v como vector propio correspondiente, para cualquier entero positivo k.
2. Si a es un valor propio de una matriz invertible A con v como vector propio correspondiente, entonces $1/a$ es un valor propio de A^{-1}, de nuevo con v como vector propio correspondiente.
3. Si a es un valor propio de A, entonces el conjunto E es un subespacio de R^n.

***Corolario* 7.1**
Sean A y S matrices de orden , entonces $A^k = A\ \Lambda^k S^{-1}$ para cada entero positivo k.

Demostración
De $S^{-1}AS = \Lambda$ obtenemos $A = S\Lambda S^{-1}$ así:

$$A^k = \underbrace{(S\Lambda S^{-1})(S\Lambda S^{-1}) \cdots (S\Lambda S^{-1})}_{K \text{ veces}}$$

Como los términos adyacentes $S^{-1}S$ se cancelan queda:

$$A^k = S\ \Lambda^k S^{-1}$$

MATRICES SIMÉTRICAS Y DIAGONALIZACIÓN ORTOGONAL

Teorema 7.8
Sea A una matriz simétrica real de , entonces los valores propios de A son reales. Para cualquier matriz simétrica, la multiplicidad algebraica de cada valor propio es igual a su multiplicidad geométrica, de modo que toda matriz simétrica es diagonalizable.

Ya vimos que vectores característicos correspondientes a valores característicos diferentes son linealmente independientes. Para matrices simétricas el resultado es más fuerte; los vectores característicos de una matriz simétrica correspondientes a valores diferentes son ortogonales.

Teorema 7.9
Sea A una matriz simétrica real de $n \times n$. Si a_1 y a_2 son valores característicos distintos con correspondientes vectores característicos reales v_1 y v_2, entonces v_1 y v_2 son ortogonales.

Demostración
Calculamos $\quad Av_1\ v_2 = a_1 v_1\ v_2 = a_1(v_1\ v_2)$
y $\quad Av_1\ v_2 = v_1\ A^t v_2 = v_1\ Av_2 = v_1\ (a_2 v_2) = a_2(v_1 v_2)$
Entonces: $\quad a_1(v_1\ v_2) = a_2(v_1\ v_2)$ como $a_1 + a_2$
Y de aquí: $\quad (v_1\ v_2) = 0$

CONTENIDO

Teorema 7.10
Sea A una matriz simétrica real de , entonces A tiene n vectores característicos reales ortonormales.

Este teorema nos dice, que si A es simétrica, entonces R^n tiene una base $B=\{u_1,u_2,\cdots,u_n\}$ de vectores característicos ortonormales de A. Sea Q la matriz cuyas columnas son $u_1,u_2,\cdots,u_n$, entonces Q es una matriz ortogonal.

> ***Definición 7.7***
> Una matriz A de se dice diagonalizable ortogonalmente, si existe una matriz ortogonal Q tal que:
>
> $$Q^tAQ = \Lambda$$
>
> Donde: $\Lambda = \text{diag.}(a_1,\cdots,a_n)$ los valores característicos de A. Como Q es ortogonal $Q^t=Q^{-1}$, lo que implica $Q^{-1}AQ = \Lambda$

Teorema 7.11
Sea A una matriz real de; entonces A es diagonalizable ortogonalmente si y solo si, A es simétrica.

Antes de ver ejemplos, enunciaremos los 3 pasos que sirven para encontrar la matriz ortogonal Q que diagonaliza a la matriz simétrica A.

1. Encuentre una base para cada espacio característico de A.
2. Encuentre una base ortonormal para cada espacio característico de A, usando el proceso de Gram-Schmidt.
3. Escriba Q como la matriz cuyas columnas son los vectores característicos ortonormales obtenidos en el paso 2.

Ejemplo 7.13
Encuentre una matriz ortogonal Q que diagonalice a la matriz A:

$$A=\begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Solución:

$$\det(A-I) = \begin{vmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{vmatrix} = (\alpha-2)^2(\alpha-8) = 8$$

Por lo que $\alpha_1=2$ y $\alpha_2=8$. Para $\alpha_1=2$, se tiene:

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Forman una base para E_2. Aplicando Gram–Schmidt se tiene:

$$v_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

El espacio característico correspondiente a $a_2 = 8$, tiene como base:

$$v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Aplicando Gram–Schmidt a v_3 se tiene:

$$v'_3 \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

Diagonaliza ortogonalmente a A

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \quad S = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

7.5. ECUACIONES EN DIFERENCIAS

En esta sección se aplica la diagonalización vista en secciones anteriores, y se verá su importancia a través de la serie de Fibonacci y las cadenas de Markov de la sección 7.6.

Las ecuaciones en diferencias se dirigen a un número finito de pasos finitos, mientras que una ecuación diferencial lleva un número infinito de pasos infinitesimales (sin embargo ambas teorías son paralelas).

Ejemplo 7.14

Supongamos que se invierten $100 por 5 años a un interés de 6%. Si se compone una vez al año, entonces el capital se multiplica por 1.06 y $P_{k+1}=1.06P_k$ donde ésta, es una ecuación en diferencias con un lapso de tiempo de un año. Relacione el capital después de k+1 años con el capital del año anterior; después de 5 años el capital original B=1000 se ha multiplicado 5 veces y:

$$P_5=(1.06)^5P_0 = (1.06)^5\ 1000 = \$\ 1338.$$

Si el lapso de tiempo se reduce a un mes:

$$P_{k+1}= (1 + 0.6/12)P_k.$$

Y después de 5 años o 60 meses:

$$P_{60}= (1 + 0.06/12)^{60}P_0= (1.005)^{60}\ 1000 = 1349$$

Si el interés se compone diariamente se tiene:

$$(1+0.06/365)^{5.365}\ 1000 = \$\ 1349.83.$$

Si se compone continuamente el interés, se añade a cada instante, entonces se tiene un proceso de límite:

$$(1+0.06/\mathrm{n})^{5n}\ 1000 \rightarrow\ e^{0.30}\ 1000 = \$\ 1349.87$$

O se puede cambiar a una ecuación diferencial, que será el límite de la ecuación en diferencias $P_{k+1}=(1+0.06\ \Delta t)P_k$, lo que implica:

$$P_{k+1} = P_k + 0.06\ \Delta t\ P_k$$
$$P_{k+1} - P_k = 0.06\ \Delta t\ P_k$$

De donde:

$$\frac{P_{k+1} - P_k}{\Delta} = 0.06\ P_k \qquad \rightarrow \qquad \frac{dp}{dt} = 0.006\ p$$

Cuya solución es $p(t) = e^{0.06t}\ p_0$,para 5 años se tiene $ 1349.07.

Este ejemplo incluyó tanto ecuaciones en diferencias como ecuaciones diferenciales, con una tendiendo a la otra a medida que desapareció el lapso de tiempo. Pero hay muchas ecuaciones en diferencias por derecho propio y aquí nuestro siguiente ejemplo:

Ejemplo 7.15 (Los conejos de Fibonacci)
Suponga que las parejas de conejos recién nacidos no tienen descendencia durante su primer mes de vida, pero que a partir de ahí cada pareja produce otra nueva pareja cada mes. Comenzando con $F_1=1$ que es la pareja recién nacida en el primer mes, hallar el número F_k de parejas en el k-ésimo mes, suponiendo que no muere ningún conejo.

En el k-ésimo mes el número de parejas de conejos es

F_k = (número de parejas más el mes anterior) +
(número de parejas recién nacidas en el k-ésimo mes)

Como nuestros conejos no tienen descendencia durante el primer mes de vida, vemos que el número de parejas recién nacidas en el k-ésimo mes es el número F_{k-2} de parejas vivas dos meses antes. Así, podemos escribir la ecuación anterior como:

$$F_k = F_{k-1} + F_{k-2}$$

Esta es una ecuación en diferencias, conocida como **Relación de Fibonacci**.

Es conveniente hacer $F_0=0$ para denotar 0 parejas en el mes 0, antes de la llegada de la primera pareja recién nacida. Así la sucesión:

$$F_0, F_1, F_2, \cdots, F_k.$$

Para el número de parejas de conejos se convierte en la sucesión de Fibonacci.

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \cdots$$

Donde cada término, comenzando con $F_2=0+1=1$ es la suma de los dos términos anteriores. Para cualquier k particular, podemos calcular F_k prolongando lo suficiente la sucesión. Sin embargo, esto puede ser una tarea tediosa; aunque solo queramos calcular F_{30}.

Una forma de hacerlo, consiste en resolver la ecuación en diferencias $F_{k+2}=F_{k+1}+F_k$; y como primer paso podemos reducirla a una ecuación $u_{k+1}=Au_k$ precisamente como el interés compuesto $P_{k+1}=1.06\ P_k$; excepto que ahora la incógnita es un vector y el multiplicador A tiene que ser una matriz:

$$\text{Si} \quad u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$$

Entonces:

$$F_{k+2} = F_{k+1} + F_k$$
$$F_{k+1} = F_{k+1}$$

Se transforma en:

$$\text{Si} \quad u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} u_k$$

Formalmente la ecuación en diferencias $u_{k+1}=Au_k$ es fácil de resolver. Como cada paso, conlleva una multiplicación por A, la solución u_k está relacionada con el valor inicial u0 mediante $u_k = A^k u_0$.; entonces el problema se reduce a encontrar A^k de un modo rápido y esto se puede resolver a través de los valores y vectores característicos de A.

Teorema 7.12

Si puede diagonalizarse A, $A = S\Lambda S^{-1}$, entonces:

$$u_k = A^k u_0 = S^k S^{-1} u_0$$

O lo que es lo mismo:

$$u_k = \begin{bmatrix} \vdots & \vdots & \vdots \\ x_1 & \cdots & x_n \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \Lambda_1^k & & \\ & \ddots & \\ & & \Lambda_n^k \end{bmatrix} S^{-1} u_0$$

Donde $S^{-1}u_0$ es un vector columna de R^n, ya que S^{-1} tiene rango n. De modo que sus columnas forman una base de R^n, de aquí que:

$$S^{-1}U_0 = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix}$$

Entonces:

$$u_k = C_1 \Lambda_1^k x_1 + \cdots + C_n \Lambda_1^k x_n$$

La solución general es una combinación de las soluciones especiales $\alpha_i^k x_i$ y los coeficientes c_i correspondientes con la condición inicial u_0 son:

$$C_1 \alpha_1^0 x_1 + \cdots + C_n \alpha_n^0 x_n = u_0 \quad \text{o} \quad Sc = u_0$$

En el caso específico de la ecuación de Fibonacci, el primer paso es diagonalizar la matriz A:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$det\,(A - \alpha I) = a^2 - a - 1$$

Lo que implica:

$$\alpha_1 = \frac{1+\sqrt{5}}{2} \qquad \alpha_{21} = \frac{1-\sqrt{5}}{2}$$

$$A = S\,\Lambda\,S^{-1} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 & \\ & \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & -\alpha_2 \\ -1 & \alpha_1 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

Como $F_0 = 0$ y $F_1 = 1$, se tiene $u_0 = [1, 0]$, entonces:

$$\begin{bmatrix} F^{k+1} \\ F_k \end{bmatrix} = u_k = A^k u_0 = S^k S^{-1} u_0 = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1^k & \\ & \alpha_2^k \end{bmatrix} \begin{bmatrix} 1 & -\alpha_2 \\ -1 & \alpha_1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1^k & \\ & \alpha_2^k \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1^k \\ -\alpha_2^k \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha_1^{k+1} & -\alpha_2^{k+1} \\ \alpha_1^k & \alpha_2^k \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

Entonces el número de Fibonacci F_k es la segunda componente de este producto:

$$F_k = \frac{\alpha_1^k}{\alpha_1 - \alpha_2} - \frac{\alpha_2^k}{\alpha_1 - \alpha_2} = 1/\sqrt{5}\left[(1+\sqrt{5}/2)^k - (1-\sqrt{5}/2)^k\right]$$

Como el segundo término $\left[1-\sqrt{5}/2\right]^k/\sqrt{5}$ siempre es menor que ½, solo debe moverse el primer término al entero más cercano, de modo que:

$$F_k \approx 1/\sqrt{5}\,(1+\sqrt{5}/2)^k$$

O bien

F_k = entero más cercano a $1/\sqrt{5}\,(1+\sqrt{5}/2)^k$ para toda k

Por ejemplo: Si k=1000
Entonces: F_{1000} = entero más cercano a $1/\sqrt{5}\ (1+\sqrt{5}/2)^{1000}$

Pero este número como se puede ver es muy grande y aún más grande es F_{1001}, la razón F_{1000}/F_{1001} debe estar muy cerca de la cantidad $1+\sqrt{5}/2 \approx 1.1618$ que los griegos llamaron la **"sección dorada"**. En otras palabras α_2^k se vuelve insignificante comparado con α_1^k y la razón F_{k+1}/F_k se acerca a $\alpha_1^{k+1}/\alpha_1^k = \alpha_1$.

7.6. CADENAS DE MARKOV

Las Cadenas de Markov se utilizaron inicialmente para analizar procesos en física y meteorología; sin embargo, las aplicaciones más recientes incluyen el análisis de los movimientos de precios de bienes, el mantenimiento de maquinaria, el comportamiento de los animales en el laboratorio, la selección de productos por el consumidor, la longitud de las colas en un supermercado o un aeropuerto, en el manejo de inventarios en cuanto a nivel y en la administración de plantas.

Para definir una cadena de Markov, considere un experimento con un espacio muestra finito $S = \{E_1, E_2, \cdots, E_n\}$ y una secuencia (o cadena) de experimentos realizados. Se dice que el experimento está en el estado E_i en el intento m-ésimo, si E_i es el resultado del m-ésimo ensayo del experimento.

Definición 7.8
Una secuencia de intentos de m experimentos es una cadena de Markov, si:

a) El resultado del intento m-ésimo depende solo del resultado del intento $(m-1)$-ésimo y no de los resultados en los intentos anteriores, y
b) La probabilidad de pasar del estado E_i al estado E_j en dos intentos sucesivos del experimento permanece constante.

Por ejemplo
Si el clima de hoy depende solamente del clima de ayer, entonces la observación y predicción del clima es un problema de cadenas de Markov. Si la probabilidad de escoger una marca particular de coche la próxima vez que piense comprar uno, depende únicamente del auto que ya se tiene, entonces el problema del patrón de ventas de automóviles es un problema que se puede resolver por cadenas de Markov.

Por otra parte, si el clima de hoy es determinado por el clima de varios días anteriores, entonces no se trata de una cadena de Markov.

Una cadena de Markov se caracteriza por las probabilidades de que el sistema pase de un estado a otro en intentos sucesivos.

MATRIZ DE TRANSICIÓN DE UNA CADENA DE MARKOV

La matriz de transición de una cadena de Markov, es la matriz de de $n \times n$ probabilidades $T = (pij)$ cuya componente ij-ésima pij es la probabilidad de que el sistema pase del estado E_i al estado E_j en intentos sucesivos del experimento; por lo cual, todos los componentes son no negativos y la suma de ellos por columna es 1.

Ejemplo 7.16

Cada año $\frac{1}{10}$ de la gente que vive fuera de California se muda dentro y $\frac{2}{10}$ de la gente que vive dentro se muda fuera.

Esto sugiere una ecuación en diferencias. Comenzamos con y_0 gente que vive fuera y z_0 que vive dentro y al final del primer año hay

$$y_1 = 0.9\, y_0 + 0.2\, z_0$$
$$z_1 = 0.1\, y_0 + 0.8\, z_0$$

O bien:

$$\begin{bmatrix} y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

La matriz T es una matriz de transición, ya que:

$$0.9+0.1=1 \qquad 0.2+0.8=1$$

Y son todos no negativos. Por lo tanto, las potencias T^k son no negativas.

Resolvemos esta ecuación en diferencias usando $S^k S^{-1} u_0$; posteriormente vemos si la población tiende a un estado estacionario. Esto es, que después de un tiempo las probabilidades de que el sistema se encuentre en cada uno de sus

posibles estados son invariables con el tiempo. Para comenzar los cálculos, T tiene que diagonalizarse.

$$T=\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$$

$$det(T-\alpha I) = \alpha_2 - 1.7\alpha + 0.7$$
$$\alpha_1 = 1 \ \ y \alpha_2 = 0.7$$

$$A= S\Lambda S^{-1} = \begin{bmatrix} ⅔ & ⅓ \\ ⅓ & ⅓ \end{bmatrix}\begin{bmatrix} 1 & \\ & 0.7 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$$

Podemos encontrar A^k y la distribución después de k años:

$$\begin{bmatrix} y_k \\ z_k \end{bmatrix} = A^k \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}\begin{bmatrix} ⅔ & ⅓ \\ ⅓ & ⅓ \end{bmatrix}\begin{bmatrix} 1 & \\ & 0.7 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

$$= y_0\, z_0 \begin{bmatrix} ⅔ \\ ⅓ \end{bmatrix} + (y_0 - 2z_0)\ (0.7)^k \begin{bmatrix} ⅓ \\ -⅓ \end{bmatrix}$$

Esta es la solución que queríamos; a largo plazo el factor (0.7)k se vuelve muy pequeño y la solución tiende a un estado límite.

$$\begin{bmatrix} y_0 \\ z_0 \end{bmatrix} = (y_0 + z_0)\begin{bmatrix} ⅔ \\ ⅓ \end{bmatrix}$$

La población total es todavía y_0+z_0, la misma que al principio, pero en el límite ⅔ de esta población está fuera de California y ⅓ está dentro. Esto es cierto independientemente de las distribuciones iniciales. Si se comienza en el año con ⅔ fuera y ⅓ dentro, entonces termina de la misma manera:

$$\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}\begin{bmatrix} ⅔ \\ ⅓ \end{bmatrix} = \begin{bmatrix} ⅔ \\ ⅓ \end{bmatrix}$$

O bien $\qquad Au_\infty = u_\infty$

El estado estacionario es el vector propio de A correspondiente a α=1. La multiplicación por A que nos lleva de un lapso de tiempo a otro no altera u_∞.
En el ejemplo de California, si el individuo está afuera, se mudará adentro con una probabilidad de 1⁄10, si está adentro, entonces se mudará afuera con probabilidad 2⁄10. Su movimiento se vuelve un proceso aleatorio y la matriz A que lo gobierna es una matriz de transición. No sabemos dónde está, pero cada año las componentes de $u_k = A^k u_0$ especifican la probabilidad de que esté fuera del estado y la probabilidad de que esté dentro. Estas probabilidades suman 1; el individuo tiene que estar en algún lado y nunca son negativas. Porque α=1 es siempre un valor propio y porque su vector propio es el estado estacionario.

Como se vio en el caso de los conejos de Fibonacci, se tiene que el vector propio o de información después del k-ésimo estado, es igual a $A^k x_1$ para un vector inicial x_1 y una matriz diagonalizable A. Como es evidente, un aspecto importante es si alguno de los valores propios de A tiene magnitud mayor que 1. Cuando éste es el caso, la magnitud de las componentes del vector de información pueden crecer exponencialmente como es el caso de la sucesión de Fibonacci donde $|a_1|>1$, se dice que es inestable.

Por otro lado, si todos los valores propios tienen magnitud menor que 1, las componentes del vector de información deben tender a cero a medida que k crece. Si las probabilidades de Markov decrecieran a cero, entonces la ecuación sería estable, cosa que no sucede porque en cada estado deben sumar 1.

Dada cualquier ecuación en diferencias $U_{k+1} = A_{uk}$, donde queremos estudiar su comportamiento cuando $k \to \infty$. Suponiendo que A puede diagonalizarse, la solución U_k será una combinación de soluciones para:

$$U_k = S\,\Lambda^k S^{-1}\,U_0 = C_{11}^k\,x_1 + \cdots + C_{nn}^k\,x_n$$

Definición 7.9
La ecuación en diferencias $U_{k+1} = A_{uk}$ es estable y $U_k \to 0$. Si todos los valores propios satisfacen $|a_i|<1$, es neutralmente estable y u_k está acotado si todo $|a_i|\leq 1$ y es inestable (u_k no está acotado) cuando al menos un valor propio de A tiene $|a_i|>1$.

7.7. FORMAS CUADRÁTICAS

Las formas cuadráticas son una aplicación inmediata de los valores propios, el teorema del eje principal asegura que toda forma cuadrática se puede diagonalizar, este resultado tiene importantes aplicaciones a vibración de cuerpos elásticos, mecánica cuántica y circuitos eléctricos.

Definición 7.10
Se dice que una ecuación de la forma

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0 \qquad (7.4)$$

Donde son números reales y al menos uno de ellos es diferente de cero (es una ecuación cuadrática en x,y). La expresión $ax^2 + 2by + cy^2$, se denomina la forma cuadrática asociada.

Ejemplo 7.17
En la ecuación cuadrática

$$3x^2 + 5xy - 7y^2 + 2x + 7 = 0$$

Las constantes son:

$a=3 \quad b=5/2 \quad c=-7 \quad d=2 \quad e=0 \quad f=7.$

Ejemplo 7.18

Ecuación cuadrática	**Forma cuadrática asociada**
$3x^2 + 5xy - 7y^2 + 2x + 7 = d$	$3x^2 + 5xy - 7y^2$
$4x^2 - 5y^2 + 8y + 9 = d$	$4x^2 - 5y^2$
$xy + y = d$	xy

Las gráficas de las ecuaciones cuadráticas en x, y se llaman **cónicas** o **secciones cónicas**. Las cónicas más importantes son la elipse, la circunferencia, la hipérbola y la parábola.

Una forma cuadrática con una variable x es un polinomio $f(x)=ax^2$con $a\neq 0$. Una forma cuadrática en 2 variables x, y es un polinomio $f(x)=ax^2+bxy+y^2$, donde a, b o c es $\neq 0$. El término cuadrático significa de grado 2, el término forma significa homogénea, es decir, cada sumando tiene un producto del mismo

número de variables, a saber 2. Se dice que una cónica tiene una posición normal relativa a los ejes de coordenadas, si no se encuentra trasladada ni rotada.

Ejemplo 7.19

La ecuación

$$\frac{x^2}{4} + \frac{y^2}{9} = 1$$

es de la forma:

$$\frac{x^2}{k^2} + \frac{y^2}{1^2} = 1 \qquad \text{con} \qquad k=2, \quad 1=3$$

Por lo tanto, su gráfica es una elipse en posición normal que interseca al eje x en:(–2,0), (2,0) y al eje y en (0,–3) y (0,3).

La ecuación $x^2–8y^2 = –16$ se puede expresar como $y^2/2 – x^2/16 = 1$ que es de la forma

$$y^2/k^2 – x^2/l^2 = 1 \qquad \text{con} \qquad k = \sqrt{2} \text{ y } l=4$$

Por lo tanto, su gráfica es una hipérbola en posición normal que corta al eje y en (0, – $\sqrt{2}$) y (0,$\sqrt{2}$).

La ecuación $5x^2+2y = 0$ se puede reescribir como: $x^2=(–2/5)\ y$, que es de la forma $x^2=ky$ con $k=–2/5$ como $k<0$. Su gráfica es una parábola en posición normal que se abre hacia abajo.

Ninguna cónica en posición normal tiene términos cruzados (xy) en su ecuación. Si éstos aparecen, significa que la cónica se giró con respecto a su posición normal. Tampoco se tienen cónicas en posición normal que tengan simultáneamente un término en x^2 y en x o en y^2 y en y. Si tales términos aparecen, significa que la cónica se trasladó de su posición normal. Ver en la figura 7.2.

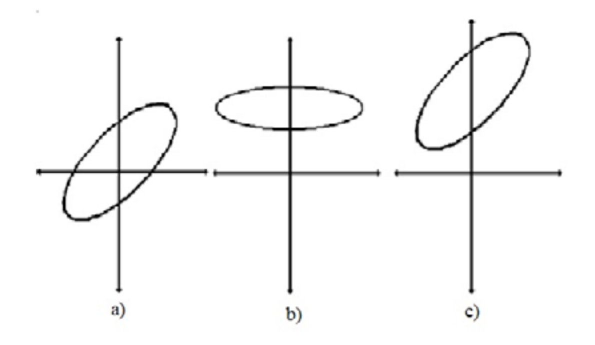


Figura 7.2 Cónica a) rotada, b) trasladada y c) rotada y trasladada

Ejemplo 7.20

Dada la ecuación cuadrática

$$2x^2 + y^2 -12x -4y + 18 = 0$$

Tiene términos en x^2, x, y^2, y; pero no en xy ; entonces, su gráfica es una cónica que se trasladó pero no se giró: Agrupamos términos para obtener:

$$(2x^2 - 12x) + (y^2 - 4y) + 18 = 0$$
$$2(x^2 - 6x) + (y^2 - 4y) = -18$$

Completando cuadrados; donde para completar el cuadrado de una expresión de la forma x^2+px se suma y se resta $(P/2)^2$.

$$x^2 + px = x^2 + px + (p/2)^2 - (p/2)^2 = (x+p/2)^2 - (p/2)^2$$

Se tiene entonces:

$$2(x^2 - 6x + 9) + (y^2 - 4y + 4) = -18 + 18 + 4$$
$$2(x - 3)^2 + (y - 2)^2 = 4$$

Si se trasladan los ejes de coordenadas mediante

$$x' = x-3 \qquad y' = y-2$$

Se tiene: $2x'^2 + y'^2 = 4$

Entonces: $\dfrac{x'^2}{2} + \dfrac{y'^2}{4} = 1$

Es una elipse en posición normal con respecto a $x'y'$ como se puede ver en la figura 7.3

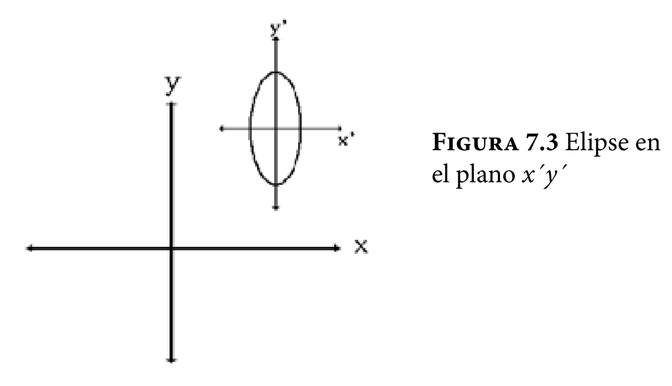


Figura 7.3 Elipse en el plano $x'y'$

CONTENIDO

La ecuación (7.4)

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

se puede expresar en forma matricial como:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + f = 0$$

Es decir: $\quad x^t Ax + kx + f = 0 \qquad (7.5)$

Donde: $\quad x = \begin{bmatrix} x \\ y \end{bmatrix} \quad A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \quad k = \begin{bmatrix} d & e \end{bmatrix}$

Con esta notación, la forma cuadrática asociada a (7.5) es:

$$x^t Ax$$

La matriz simétrica A se llama la matriz de la forma cuadrática $x^t Ax$.

Ejemplo 7.21

La matriz de la forma cuadrática

$$3x^2 + 5xy - 7y^2$$

es: $\begin{bmatrix} 3 & 5/2 \\ 5/2 & -7 \end{bmatrix}$

Considere que la ecuación de una cónica C, es:

$$x^tAx + kx + f = 0 \qquad (7.6)$$

A continuación, se muestra que siempre es posible rotar los ejes de coordenadas x–y de tal manera, que la ecuación de la cónica con respecto al sistema de coordenadas x'–y' no contenga un término en xy.

Etapa 1
Encuentre la matriz

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix}$$

Que diagonalice ortogonalmente a A.

Etapa 2
Si es necesario, intercambiar las columnas de P de tal suerte que det (P)=1. Esto garantiza que la transformación ortogonal de coordenadas (7.7) es una rotación:

$$x = Px' \qquad (7.7)$$

Etapa 3
Para obtener la ecuación de C con respecto al sistema $x'\,y'$, se sustituye (7.6) en (7.7)

$$(Px')^t \; A(Px') + k(Px') + f = 0$$

o

$$x^t\,(P^tAP)\,x' + (kP)x' + f = 0 \qquad (7.8)$$

Como P diagonaliza ortogonalmente a A entonces

$$P'AP = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}$$

Donde α_1 y α_2 son valores propios de A, por lo cual (7.8) se puede expresar como:

$$[x' \quad y']\begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} + [\,d \quad C\,]\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} f = 0$$

Entonces: $\alpha_1 x'^2 + \alpha_2 y'^2 + d'x' + e'y' + f = 0.$

Donde: $d' = dP_{11} + eP_{21} \quad e' = dP_{12} + eP_{22}.$

El siguiente teorema resume todo el proceso.

Teorema 7.13: De los ejes principales para R²

Sea $ax^2 + 2bxy + cy^2 + dx + cy + f = 0$

la ecuación de una cónica C y sea

$$x^t Ax = ax^2 + 2bxy + cy^2$$

la forma cuadrática asociada. Entonces los ejes de coordenadas se pueden rotar de tal manera que la ecuación C con respecto al nuevo sistema de coordenadas $x'\,y'$ tiene la forma:

$$\alpha_1 x'^2 + \alpha_2 y'^2 + d'x' + e'y' + f = 0$$

Donde α_1 y α_2 son valores propios de A. La rotación se puede efectuar mediante la sustitución $x = Px'$, donde P diagonaliza ortogonalmente a A.

Ejemplo 7.22

Describa la cónica C cuya ecuación está dada por:

$$5x^2 - 4xy + 8y^2 + (20/\sqrt{5})x - (80/\sqrt{5})y + 4 = 0$$

La forma matricial de esta ecuación es:

$$x^t Ax + kx + 4 = 0$$

Donde: $A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}$ $k = [20/\sqrt{5}, -80/\sqrt{5}]$

$$det\,(A - \alpha I) = (\alpha - 9)(\alpha - 4)$$

Y v_1, v_2 son vectores columna de la matriz P como sigue:

$$P = \begin{bmatrix} 2/5 & -1/\sqrt{5} \\ 1/5 & -2/\sqrt{5} \end{bmatrix}$$

P diagonaliza ortogonalmente a A. Sustituyendo $x = Px'$

$$(Px')^t A(Px') + k(Px') + y = 0$$

Es decir: $(x'^t)(P^t AP)x' + (kP)x' + y = 0$

Dado que

$$P^t AP = \begin{bmatrix} 4 & \\ & 9 \end{bmatrix};\ kP = [20/\sqrt{5}, -80/\sqrt{5}] \begin{bmatrix} 2/5 & -1/5 \\ 1/5 & 2/5 \end{bmatrix} = [-8, -36]$$

Entonces: $4x'^2 + 9y'^2 - 8x' - 36y' + 4 = 0$

Trasladando: $4(x'^2 - 2x') + 9(y'^2 - 4y') = -4$

Completando cuadrados: $4(x'^2 - 2x' + 1) + 9(y'^2 - 4y' + 4) = -4+4+36$

O bien: $4(x'-1)^2 + 9(y'-2)^2 = 36$

Sean $x'' = x' - 1$
$y'' = y' - 2$

las ecuaciones de traslación, entonces: $4x''^2 + 9y''^{12} = 36$

o $\dfrac{x''^2}{9} + \dfrac{y''^2}{4} = 1$ es la ecuación de una elipse.

Ejemplo 7.23

Describa la cónica C cuya ecuación está dada por:

$$5x^2 - 4xy + 8y^2 - 36 = 0$$

O bien $\qquad x^tAx - 36 = 0$

Donde: $\qquad A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}$

$$det\ (\ A - \alpha I) = (\alpha - 9)\ (\alpha - 4) = 0$$

Para α=4, se tiene:

$$v_1 = \begin{bmatrix} ^{2}/\sqrt{5} \\ ^{1}/\sqrt{5} \end{bmatrix}$$

Para α=9, se tiene:

$$v_2 = \begin{bmatrix} ^{-1}/\sqrt{5} \\ ^{2}/\sqrt{5} \end{bmatrix}$$

Por lo tanto:

$$P = \begin{bmatrix} \frac{2}{5} & -1\sqrt{5} \\ ^{-1}/\sqrt{5} & ^{2}/\sqrt{5} \end{bmatrix}$$

Diagonaliza ortogonalmente a A; además el det P=1; por lo cual, se puede concluir que x=Px' es una rotación o bien:

$$(x')^t\ (P^tAP)\ x' - 36 = 0$$

Donde:

$$P'\ PA = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$$

$$[x' \quad y'] \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} - 36 = 0$$

$$4x'^2 + 9y'^2 - 36 = 0$$

o $\dfrac{x'^2}{9} + \dfrac{y'^2}{4} = 1$ es una elipse.

7.8. RESUMEN

1. Si A es una matriz de, entonces un escalar α es un valor propio de A si hay un vector v distinto de cero; tal que:

   $Av = \alpha v$

   El vector $v \neq 0$ se llama un vector propio de A correspondiente a un valor propio α

2. Si A es de orden $n \times n$, entonces la ecuación

   $(A - \alpha I)v = 0$

   Es un sistema homogéneo de n ecuaciones con n incógnitas, como por hipótesis (ya que v es un vector propio distinto de cero) el sistema tiene soluciones no triviales, entonces:

   $det\ (A - \alpha I) = 0$

   A esta ecuación se le denomina la ecuación característica de A.

3. Si contamos multiplicidades, cada matriz de tiene exactamente n valores característicos.

4. Si A es una matriz de orden, entonces las siguientes afirmaciones son equivalentes:
   a) a es un valor característico de A.

b) El sistema $(A-\alpha I)v = 0$ tiene soluciones no triviales.
c) Existe un vector $v \neq 0$ en R^n tal que $Av = av$

5. Si a es un valor característico de A, entonces el espacio solución del sistema de ecuaciones $(A-\alpha I)v = 0$ se denomina el espacio característico de A correspondiente a α, y los vectores diferentes de cero en el espacio característico, se denominan los vectores característicos de A correspondientes a a.

6. Sea A una matriz de y sean $a_1, a_2, \cdots, a_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $v_1, v_2, \cdots, v_m$. Entonces v_i son linealmente independientes o lo que es equivalente, los vectores característicos correspondientes a valores característicos diferentes son linealmente independientes. (La demostración se hace por inducción).

7. Sea a un valor característico de A, entonces la multiplicidad geométrica de α es la dimensión del espacio característico correspondiente a α (que es la nulidad de la matriz $A-\alpha I$)s, esto es:

   Multiplicidad geométrica de $a = \dim E_a$

8. Sea A una matriz de orden $n \times n$, entonces A tiene n vectores característicos linealmente independientes si y solo si la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, A tiene n vectores característicos linealmente independientes si todos sus valores característicos son distintos (pues si no, podrían ser menores que n).

9. Sea la matriz A de orden $n \times n$.
   a) Si α es un valor propio de A con v como vector propio correspondiente, entonces α^k es un valor propio de A^k, de nuevo con v como vector propio correspondiente, para cualquier entero positivo k.
   b) Si a es un valor propio de una matriz invertible A con v como vector propio correspondiente, entonces $1/\alpha$ es un valor propio de A^{-1}, de nuevo con v como vector propio correspondiente.
   c) Si α es un valor propio de A, entonces el conjunto E_α es un subespacio de R^n.

10. Una secuencia de intentos de m experimentos es una cadena de Markov si:
   a) El resultado del intento m-ésimo, depende solo del resultado del intento (m–1)-ésimo y no de los resultados en los intentos anteriores
   b) La probabilidad de pasar del estado E_i al estado E_j en dos intentos sucesivos del experimento permanece constante.

11. Se dice que una ecuación de la forma

$$\alpha x^2 + 2bxy + cy^2 + dx + cy + f = 0$$

Donde son números reales y al menos uno de ellos es diferente de cero, es una ecuación cuadrática en x,y. La expresión $ax^2 + 2by + cy^2$, se denomina la forma cuadrática asociada.

12. Sea $\alpha x^2 + 2bxy + cy^2 + dx + cy + f = 0$
la ecuación de una cónica C y sea

$$x^t Ax = \alpha x^2 + 2bxy + cy^2$$

la forma cuadrática asociada. Entonces los ejes de coordenadas se pueden rotar de tal manera que la ecuación C con respecto al nuevo sistema de coordenadas $x'\, y'$ tiene la forma:

$$\alpha_1 x^{12} + \alpha_2 y^{12} + d'x' + e'y' + f = 0$$

Donde α_1 y α_2 son valores propios de A. La rotación se puede efectuar mediante la sustitución $x = Px'$, donde P diagonaliza ortogonalmente a A.

7.9. APLICACIONES DE VALORES Y VECTORES PROPIOS

(Tomado de *Aplicaciones de álgebra lineal*, Rorres y Anton, 1977)

En esta sección se presentan aplicaciones a la demografía además de las aplicaciones vistas en secciones anteriores, cabe recordar que este tema cuenta con una gran cantidad de usos y aplicaciones en diferentes problemas de la ingeniería y de otras áreas del conocimiento.

CRECIMIENTO DE LA POBLACIÓN Y LAS EDADES ESPECÍFICAS

Uno de los modelos que más comúnmente utilizan los demógrafos para el crecimiento de la población es conocido como "modelo de Leslie", desarrollado en los años cuarenta. Este modelo describe el crecimiento de la parte femenina de una población, humana o animal, clasificando a las hembras por edades en clases o intervalos de igual número. Especificando, suponga que la edad máxima alcanzada por una hembra cualquiera de la población en L años (o alguna otra unidad de tiempo) y que la población se divide en n clases de edades. Cada clase tiene L/n años de duración. La identificación de cada clase y los intervalos correspondientes, se harán de acuerdo con la siguiente tabla:

Clases de edades	Intervalos de edades
1	$[0, L/n)$
2	$[L/n, 2L/n)$
3	$[2L/n, 3L/n)$
⋮	⋮
n–1	$[(n-2)L/n,. (n-1)L/n)$
n	$[(n-1)L/n, L]$

Suponga conocido el número de hembras de casa una de las n clases en el tiempo t=0. En particular, suponga que hay $x_1^{(0)}$ hembras en la primera clase, $x_2^{(0)}$ en la segunda y así sucesivamente. Con estos n números se puede formar el vector columna $x^{(0)}$ siguiente:

$$x^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix}$$

A este vector es al que se le llama *vector de la distribución inicial de edades.*

A medida que transcurre el tiempo y debido a los procesos biológicos (nacimiento, envejecimiento y muerte), cambia el número de hembras que hay en cada una de las n clases. Describiendo los procesos biológicos cuantitativamente,

se verá que el vector de la distribución inicial de edades puede proyectarse hacia el futuro.

La forma más fácil de estudiar el proceso del envejecimiento consiste en hacer observaciones directas de la población en tiempos discretos, como por ejemplo 1_0, 1_1, 1_2, $\cdots$ 1_k, $\cdots$. El modelo de Leslie requiere que la duración entre dos tiempos sucesivos de observación sea igual a la duración de los intervalos de edad. Por lo tanto, se tiene que

$$
\begin{aligned}
t_0 &= 0 \\
t_1 &= \frac{L}{n} \\
t_2 &= \frac{2L}{n} \\
&\vdots \\
t_k &= \frac{kL}{n} \\
&\vdots
\end{aligned}
$$

Con esta suposición, todas las hembras de orden $(i+1)$ en el tiempo t_{k+1}, estaban en la clase de orden i en el tiempo t_k.

Los procesos de nacimiento y muerte entre dos tiempos sucesivos de observación se pueden por medio de los siguientes parámetros demográficos:

a_i $i = 1, 2, \cdots, n$	Es promedio del número de hijas que tiene una hembra durante el tiempo que permanece en la clase de orden i
b_i $i = 1, 2, \cdots, n-1$	Es fracción de las hembras de la clase de orden que se espera sobrevivan y pasen a la clase de orden $i+1$

Por definición, se tiene que

(i) $a_i \geq 0$ para $i = 1, 2, \cdots, n$
(ii) $0 < b_i < 1$ para $i = 1, 2, \cdots, n-1$

Note que ningún puede ser igual a cero, ya que de ocurrir esto, ninguna hembra sobrevivirá más allá de la clase de orden i. También se supone que hay por lo menos un valor de se le llama *clase fértil.*

En seguida el vector de la distribución de las edades, $x^{(k)}$, en un tiempo t_k,

$$x^{(k)} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix}$$

Expresión en la cual $x_i^{(k)}$ es el número de hembras de la clase de orden i en el tiempo t_k. Por otra parte, en el tiempo t_k, las hembras de la primera clase son únicamente, las nacidas entre los tiempos t_{k-1} y t_k. En consecuencia, puede escribirse

$$\left\{\begin{matrix}\text{El número de hembras} \\ \text{de la clase 1 en el tiempo } t_k\end{matrix}\right\} = \left\{\begin{matrix}\text{Al número de hijas} \\ \text{clase 1 entre los tiempos} \\ \text{clase 1 en los tiempos} \\ t_{k-1}\ y\ t_k\end{matrix}\right\}$$

$$+\left\{\begin{matrix}\text{El número de hijas} \\ \text{de las hembras} \\ \text{de la clase 2 entre} \\ \text{los tiempos } t_{k-1}\ y\ t_k\end{matrix}\right\} + \cdots + \left\{\begin{matrix}\text{El número de hijas} \\ \text{de las hembras} \\ \text{de la clase } n \text{ entre} \\ \text{los tiempos } t_{k-1}\ y\ t_k\end{matrix}\right\}$$

O bien, matemáticamente,

$$x_1^{(k)} = a_1 x_1^{(k-1)} + a_2 x_2^{(k-1)} + \cdots + a_n x_n^{(k-1)} \qquad (7.9)$$

El número de hembras de la clase de orden $(i+1)$, donde $i=1, 2, \cdots, n-1$, en el tiempo t_{k-1} que todavía están vivas en el tiempo t_k. Entonces

$$\left\{\begin{matrix}\text{El número} \\ \text{de hembras de la clase} \\ i+1 \text{ en el tiempo } t_k\end{matrix}\right\} = \left\{\begin{matrix}\text{La fracción de hembras} \\ \text{de la clase } i \text{ que} \\ \text{sobreviven y pasan a la} \\ \text{clase } i+1\end{matrix}\right\}$$

$$\left\{ \begin{array}{l} \text{El número de hembras} \\ \text{de la clase } i \text{ tiempo } t_{k-1} \end{array} \right\}$$

O bien, matemáticamente,

$$x_{1+1}{}^{(k)} = b_i\, x_i^{(k-1)},\ i{=}1, 2, \cdots, n{-}1 \tag{7.10}$$

Usando la notación matricial, las ecuaciones (7.9) y (7.10) pueden escribirse en la siguiente forma

$$\begin{bmatrix} x_1{}^{(k)} \\ x_2{}^{(k)} \\ x_3{}^{(k)} \\ \vdots \\ x_n{}^{(k)} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{n-1} & a_n \\ b_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & b_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{n-1} & 0 \end{bmatrix} \begin{bmatrix} x_1{}^{(k-1)} \\ x_2{}^{(k-1)} \\ x_3{}^{(k-1)} \\ \vdots \\ x_n{}^{(k-1)} \end{bmatrix}$$

En forma más compacta

$$x^{(k)} = L_x{}^{(k-1)},\ k{=}\ 1, 2, \cdots \tag{7.11}$$

Donde L es la *matriz de Leslie.*

$$L{=} \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{n-1} & a_n \\ b_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & b_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{n-1} & 0 \end{bmatrix} \tag{7.12}$$

De la ecuación (7.11), se tiene que

$$\begin{aligned} x^{(1)} &{=} Lx^{(0)} \\ x^{(2)} &{=} Lx^{(1)} = L^2x^{(0)} \\ x^{(3)} &{=} Lx^{(2)} = L^3x^{(0)} \\ x^{(k)} &{=} Lx^{(k-1)} = L^kx^{(0)} \end{aligned} \tag{7.13}$$

Así, la distribución inicial de las edades, $x^{(0)}$, y la matriz de Leslie, L, son conocidas, se puede determinar de la distribución de las edades de las hembras en cualquier tiempo futuro.

Ejemplo 7.24

Suponga que la edad máxima alcanzada por las hembras de una población animal es de 15 años y que esta población se divide en tres clases de edades iguales, con intervalos de cinco años. Suponga que la matriz de Leslie para esta población es

$$L=\begin{bmatrix} 0 & 4 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \end{bmatrix}$$

Si inicialmente hay 1000 hembras en casa una de las clases, aplicando la ecuación (9.3), se tiene

$$L=\begin{bmatrix} 1{,}000 \\ 1{,}000 \\ 1{,}000 \end{bmatrix}$$

$$x^{(1)} = Lx^{(0)} = \begin{bmatrix} 0 & 4 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \end{bmatrix}\begin{bmatrix} 1{,}000 \\ 1{,}000 \\ 1{,}000 \end{bmatrix} = \begin{bmatrix} 7{,}000 \\ 500 \\ 250 \end{bmatrix}$$

$$x^{(2)} = Lx^{(1)} = \begin{bmatrix} 0 & 4 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \end{bmatrix}\begin{bmatrix} 7{,}000 \\ 500 \\ 250 \end{bmatrix} = \begin{bmatrix} 7{,}000 \\ 3{,}500 \\ 125 \end{bmatrix}$$

$$x^{(3)} = Lx^{(2)} = \begin{bmatrix} 0 & 4 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \end{bmatrix}\begin{bmatrix} 2{,}750 \\ 3{,}500 \\ 125 \end{bmatrix} = \begin{bmatrix} 4{,}375 \\ 1{,}375 \\ 875 \end{bmatrix}$$

Por lo que, después de 15 años, habrá 14,375 hembras entre o y 5 años de edad, 1,375 entre 5 y 10, y 875 entre 10 y 15 años de edad.

COMPORTAMIENTO EN EL LÍMITE

Aunque la ecuación (7.13) de la distribución de las edades de la población en cualquier tiempo, no aporta, de inmediato, un cuadro general de la dinámica

del proceso de crecimiento. Por esto, es necesario investigar qué ocurre con los valores propios y los vectores propios de la matriz de Leslie. Los valores propios de L son las raíces del polinomio característico correspondiente.

$$p(\lambda) = |\lambda \boldsymbol{I} - \boldsymbol{L}|$$

$$= \lambda^n - a_1\lambda^{n-1} - a_2 b_1 \lambda^{n-2} - a_3 b_1 b_2 \lambda^{n-3} - \cdots - a_n b_1 b_2 \cdots b_{n-1}$$

Para analizar sus raíces es conveniente introducir la función

$$p(\lambda) = \frac{a_1}{\lambda} + \frac{a_2 b_2}{\lambda^2} + \frac{a_n b_1 b_2}{\lambda^3} + \cdots + \frac{a_n b_1 b_2}{\lambda^n} \cdots \; b_{n-1} \tag{7.14}$$

Con ella, la ecuación característica puede escribirse de la forma (verificar)

$$\text{p}(\lambda) = 1 \text{ para } \lambda \neq 0. \tag{7.15}$$

Como ningún valor de o es negativo, decrece en forma monótona para los valores de mayores que cero. Además, tiene una asíntota vertical en y tiene a cero cuando . Por lo tanto, como se indica en la figura 7.4, hay un valor único de , por ejemplo, para el cual . Es decir, la matriz L tiene un valor propio único positivo. Se puede también demostrar que es simple, es decir, que tiene multiplicidad uno. Aunque aquí se omitieran los detalles del cálculo, el lector puede verificar que un vector propio de , es decir, una solución vectorial diferente de cero de

$$Lx_1 = \lambda_1 x_1$$

es

$$x_1 = \begin{bmatrix} 1 \\ b_1 / \lambda_1 \\ b_1\, b_2 / \lambda_1{}^2 \\ b_1\, b_2\, b_3 / \lambda_1{}^3 \\ b_1\, b_2\, b_3 / \lambda_1{}^3 \\ \vdots \\ b_1\, b_2 \cdots b_{n-1} / \lambda_1{}^{\text{n}-1} \end{bmatrix} \tag{7.16}$$

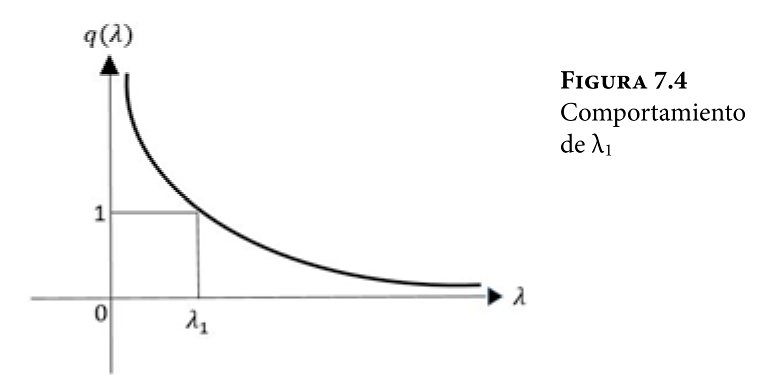


Figura 7.4
Comportamiento de λ_1

Como λ_1 es simple, el espacio propio correspondiente es unidimensional, por lo tanto, cualquier vector propio de λ_1 es un múltiplo de x_1. Estos resultados se resumen en el siguiente teorema:

Teorema 7.14

Una matriz de Leslie L, tiene un valor propio positivo único λ_1. Este valor propio es simple y tiene un vector propio x_1 cuyas entradas son todas positivas.

Ahora, se demostrará que el comportamiento, a largo plazo, de la distribución de las edades de la población, queda determinado por el valor propio positivo λ_1 y el vector propio correspondiente x_1.

Teorema 7.15

Si λ_1 es el valor propio único positivo de una matriz de Leslie L y si λ_1 es cualquier otro valor propio real o completo de L, entonces, $|\lambda_1| \leq \lambda_1$.

Por el teorema 7.15, a se le conoce como *valor propio dominante* de L. Para el propósito de este capítulo se requiere que $|\lambda_1| \leq \lambda_1$. Para todos los valores propios de L. Si este es el caso, se dice que λ_1 es un *valor propio estrictamente dominante* de L. Sin embargo, como se comprueba en el siguiente ejemplo, no todas las matrices de Leslie satisfacen esta condición.

Ejemplo 7.25

Sea

$$L=\begin{bmatrix} 0 & 0 & 6 \\ ½ & 0 & 0 \\ 0 & ⅓ & 0 \end{bmatrix}$$

El polinomio característico de L es

$$p(\lambda)=|\lambda\, I-L|=\lambda^3-1.$$

Y por lo tanto, los valores propios de L son las soluciones de $\lambda^3\times 1$

$$\lambda=1,-\frac{1}{2}+\frac{\sqrt{3}}{2}i,-\frac{1}{2}-\frac{\sqrt{3}}{2}i.$$

El valor absoluto de los tres valores propios es uno, por lo tanto, el valor propio único positivo $\lambda_1=1$ no es estrictamente dominante. Note que la matriz tiene la propiedad de que $L^3=I$. Esto significa que para cualquiera que sea la distribución inicial de edades, $x^{(0)}$, se tiene

$$X^{(0)}=X^{(3)}=X^{(6)}=\cdots=X^{(3K)}=\cdots.$$

Así, el vector de la distribución de edades oscila con periodo de tres unidades de tiempo. Tales oscilaciones (u *ondulaciones de la población,* como se les llama) no podrían ocurrir si λ_1 fuera estrictamente dominante, como se verá posteriormente.

El análisis de las condiciones necesarias y suficientes para que λ_1 sea un valor propio estrictamente dominante, queda fuera del alcance de este libro. Sin embargo, se enunciará, sin demostración, la siguiente condición suficiente:

Teorema 7.16

Si dos entradas sucesivas a_i y a_{i+1}, de la primera fila de una matriz de Leslie L son diferentes de cero, el valor propio positivo de L es estrictamente dominante.

Por lo tanto, si la población femenina tiene dos clases de edad fértil sucesivas, la matriz de Leslie correspondiente tiene un valor propio estrictamente dominante. Esto sucede siempre en la realidad si los intervalos para las poblaciones

realistas de edad de las clases de la población son lo suficientemente pequeños. Note que en el ejemplo 7.25, solo hay una clase que es fértil (la tercera) y por lo tanto no se satisface la condición del teorema 7.16.

Suponga que se puede diagonalizar la matriz L. Esto realmente no es necesario para las conclusiones a las que se pretende llegar, pero simplifica los argumentos. En este caso L tiene valores propios, $\lambda_1, \lambda_2, \cdots, \lambda_n$, no necesariamente distintos y n vectores propios linealmente independientes, $x_1, x_2, \cdots, x_n$, correspondientes. En este listado se coloca primero al valor propio que es el estrictamente dominante. Se construirá una matriz P cuyas columnas serán los valores propios de L:

$$p=[x_1 \vdots x_2 \vdots x_3 \vdots \cdots \vdots x_n]$$

Entonces, la diagonalización de L queda expresada por la ecuación

$$L = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} p^{-1}$$

De la cual se tiene que

$$L^k = P\begin{bmatrix} \lambda_1{}^k & 0 & \cdots & 0 \\ 0 & \lambda_2{}^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n{}^k \end{bmatrix} p^{-1}$$

Para k=1, 2, ⋯. Entonces, para cualquier vector de la distribución inicial de edades, $x^{(0)}$, se tiene

$$L^k x^{(0)} = PL^k = P\begin{bmatrix} \lambda_1{}^k & 0 & \cdots & 0 \\ 0 & \lambda_2{}^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n{}^k \end{bmatrix} p^{-1}x^{(0)}$$

Para k=1, 2, ⋯. Dividiendo los dos miembros de esta ecuación por λ_1^k y sabiendo que $x^{(k)}=L^k x^{(0)}$, se tiene

$$\frac{1}{\lambda_1^k} x^{(0)} = P \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & (\lambda_2/\lambda_1)^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & (\lambda_2/\lambda_1)^k \end{bmatrix} p^{-1} x^{(0)} \tag{7.17}$$

Como λ_1 es el valor propio estrictamente dominante, $|\lambda_i/\lambda_1| < 1$ para i=2, 3, ⋯, n se concluye que

$$(\lambda_2/\lambda_1)^k \rightarrow 0 \text{ cuando } k \rightarrow \infty \text{ para } i = 2, 3, \cdots,$$

Ahora, puede tomarse el límite de los dos miembros de la ecuación (7.17) para obtener

$$\lim_{k\rightarrow\infty} \frac{1}{\lambda_1^k} \boldsymbol{x}^{(k)} = P \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} p^{-1} x^{(0)} \tag{7.18}$$

$$\lim_{k\rightarrow\infty} \left(\frac{1}{\lambda_1^k} \boldsymbol{x}^{(k)} \right) = c\boldsymbol{x}_1 \tag{7.19}$$

La ecuación (7.19) da la aproximación

$$x^{(k)} \approxeq c\lambda_1^k \boldsymbol{x}_1 \tag{7.20}$$

Para valores grandes de k. Por la ecuación (7.20) se tiene también que

$$x^{(k-1)} \approxeq c\lambda_1^k \boldsymbol{x}_1 \tag{7.21}$$

Comprobando las ecuaciones (7.20) y (7.21), se ve que

$$x^{(k)} \approxeq \lambda_1^k \boldsymbol{x}^{(k-1)} \tag{7.22}$$

para valores grandes de k. Esto significa que, para valores grandes del tiempo, ceda vector de la distribución de edades es un múltiplo escalar del vector de la distribución de edades inmediato anterior, siento el escalar el valor propio positivo de la matriz de Leslie. Entonces, la *proporción* de hembras en casa una de las clases se vuelve constante. Como se verá con el siguiente ejemplo, estas proporciones en el límite pueden determinarse a partir del vector propio x_1.

Ejemplo 7.26 (ya presentado con anterioridad)
La matriz de Leslie en el ejemplo 7.24 era

$$L=\begin{bmatrix} 0 & 4 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \end{bmatrix}$$

El polinomio característico es $P(\lambda) = \lambda^3 - 2\lambda - \frac{3}{8}$ y el lector puede verificar que el valor propio positivo es $\lambda_1 = \frac{3}{8}$. Por (7.16), el vector propio correspondiente x_1 es

$$x_1=\begin{bmatrix} 1 \\ b_1/\lambda_1 \\ b_1 b_2 / \lambda_1^2 \end{bmatrix} = \begin{bmatrix} 1 \\ (\frac{1}{2}) / (\frac{3}{2}) \\ (\frac{1}{2})(\frac{1}{4})(\frac{3}{2})^2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{3} \\ \frac{1}{18} \end{bmatrix}$$

Por (7.22) se tiene $\boldsymbol{x}^{(k)} \approxeq (\frac{3}{2})\, \boldsymbol{x}^{(k-1)}$

Para valores grandes de k. Por lo tanto, cada cinco años, aumentará en aproximadamente un 50% el número de hembras en cada una de las tres clases, como aumentará también el número total de hembras de la población.

Por (7.20) se tiene $\boldsymbol{x}^{(k)} \approxeq c\, (\frac{3}{2})\, \boldsymbol{x}^{k} \begin{bmatrix} 1 \\ \frac{1}{3} \\ \frac{1}{18} \end{bmatrix}$

En consecuencia, las hembras estarán distribuidas, finalmente, de acuerdo con las relaciones $1 : \frac{1}{3} : \frac{1}{18}$. Esto corresponde a una distribución del 72% en las hembras en la primera clase, 24% en la segunda y 4% en la tercera.

Ejemplo 7.27

Para este ejemplo se utilizarán los parámetros de nacimientos defunciones de las mujeres canadienses a partir del año de 1965. Como son pocas las mujeres que se embarazan después de los 50 años, el problema se limitará a la parte de la población femenina cuya edad está comprendida entre 0 y 50 años. Los datos presentan clases de 5 años, por lo que se tendrá un total de diez clases de edades. En lugar de escribir toda la matriz de Leslie, 10×10, se presentará una lista de los parámetros de nacimientos y defunciones, como sigue:

Intervalos de edades	a_i	b_i
[0,5)	0.00000	0.99651
[5, 10)	0.00024	0.99820
[10, 15)	0.05861	0.99802
[15,20)	0.28608	0.99729
[20, 25)	0.44791	0.99694
[25,30)	0.36399	0.99621
[30,35)	0.22259	0.99460
[35,40)	0.10457	0.99184
[40,45)	0.02826	0.98700
[45, 50)	0.00240	–

Por medio de las técnicas numéricas, el valor propio positivo y el vector propio correspondiente son

$$\lambda_1 = 1.07622 \text{ y } x_1 = \begin{bmatrix} 1.00000 \\ 0.92594 \\ 0.85881 \\ 0.797641 \\ 0.73800 \\ 0.68364 \\ 0.63281 \\ 0.58482 \\ 0.53897 \\ 0.49429 \end{bmatrix}$$

Así, si las mujeres canadienses continuaran respondiéndose y muriendo de la misma forma que en 1065, finalmente cada cinco años, los números aumentarán en 7.622%. Por el vector propio x_i se ve que, el límite, por cada 100,00 hembras

CONTENIDO

cuya edad esté comprendida entre 0 y 5 años, habrá 92,594 hembras cuya edad esté comprendida entre los 5 y 10 años, 85.881 entre los 10 y los 15 y así sucesivamente.

Considere de nuevo la ecuación (7.20) que da el vector de la distribución de la población por edades, para valores grandes del tiempo:

$$\boldsymbol{x}^{(\mathrm{k})} \approxeq c\lambda^{1k}\, x_1 \tag{7.23}$$

Se presentan tres casos que dependen del valor del valor propio positivo,

i. La población finalmente crece si $\lambda_1 > 1$
ii. La población finalmente decrece si $\lambda_1 < 1$
iii. La población se estabiliza si $\lambda_1 = 1$.

El caso $\lambda_1 = 1$ es de particular interés porque determina una *población de crecimiento cero.* Para cualquier distribución inicial de las edades, la población tiene a una distribución de las edades en el límite que es algún múltiplo del vector propio. Por las ecuaciones (7.14) y (7.15) se ve que $\lambda_1 = 1$ es un valor propio si, y solo si,

$$a_1 + a_2b_1 + a_3b_1b_2 + \cdots + a_nb_1b_2 \cdots b_{n-1} = 1 \tag{7.24}$$

La expresión

$$R = a_1 + a_2b_1 + a_3b_1b_2 + \cdots + a_nb_1b_2 \cdots b_{n-1} \tag{7.25}$$

se llama *tasa neta de la reproducción* de la población. Entonces, se puede decir que una población es de crecimiento nulo si, y solo si, su tasa neta de la reproducción es igual a la unidad.

7.10. NOTAS HISTÓRICAS



La primera aparición de los **valores propios** fue en relación con su uso para resolver ecuaciones diferenciales. En 1743, **Leonhard Euler (Basilea, Suiza, 15 de abril de 1707 – San Petersburgo, Imperio ruso, 18 de septiembre de 1783)** introdujo por primera vez el método estándar de resolución de una ecuación diferencial de n-ésimo orden con coeficientes constantes

$$y^{(n)} + a_{n-1}\, y^{(n-1)} +\cdots+ a_1\, y^1 + a_0 y = 0$$

Usando funciones de la forma $y{=}e^{at}$, donde a es una raíz de la ecuación característica:

$$z^n + a_{n-1}\, z^{n-1} +\cdots+ a_1\, z + a_0 = 0$$

Esta es la misma ecuación que se obtiene al hacer las sustituciones

$$y_1 = y,\, y_2 = y'\,,\, y_3 = y",..,\, y_n = y^{(n-1)},$$

Reemplazando la única ecuación de orden n por un sistema de n ecuaciones de primer orden

$$\left|\begin{array}{lcl} y'_1 & = & y_2 \\ y'_2 & = & y_3 \\ y'_n & = & -a_0\, y_1 - a\;\; y_2 - \cdots - a_n - 1 y_n \end{array}\right.$$

Y calculando la ecuación característica de la matriz de coeficientes de este sistema.

David Hilbert

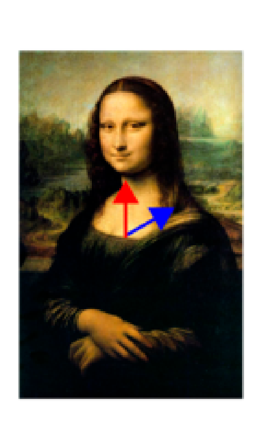


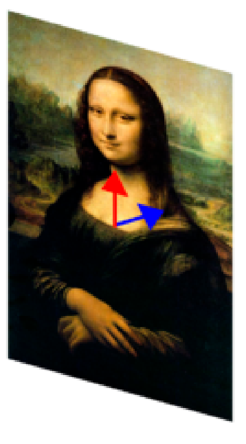
23 enero 1862, Königsberg (Prusia)-14 febrero 1943, Göttingen (Alemania).

El nombre de Hilbert ocupa un lugar muy especial en el imaginario colectivo de los matemáticos. Sin duda se trata del

matemático más famoso del siglo XX, a lo que contribuyeron de manera muy especial su aportación a la configuración de los métodos axiomáticos actuales, sus profundos resultados en álgebra, teoría de números, geometría y teoría de funciones, los celebérrimos "problemas matemáticos" que dejó planteados en 1900, y las venturas y desventuras de sus intentos de resolver la cuestión de los fundamentos de la matemática. En el año de su muerte, se le celebraba como aquel "a quien el mundo consideró durante las últimas décadas como el más grande matemático vivo".

La palabra alemana Eigen, que se traduce en español como propio, se usó por primera vez en este contexto por David Hilbert en 1904 (aunque Helmholtz la usó previamente con un significado parecido). Eigen se ha traducido también como inherente, característico o el prefijo auto–, donde se aprecia el énfasis en la importancia de los valores propios para definir la naturaleza única de una determinada transformación lineal. Las denominaciones: vector y valor característico también se utilizan habitualmente. (Fuente: Wikipedia)





Fuente: https://es.wikipedia.org/wiki/Vector_propio_y_valor_propio

En esta transformación de la Mona Lisa, la imagen se ha deformado de tal forma que su eje vertical no ha cambiado. (nota: se han recortado las esquinas en la imagen de la derecha). El vector azul, representado por la flecha azul que va desde el pecho hasta el hombro, ha cambiado de dirección, mientras que el rojo, representado por la flecha roja, no ha cambiado. El vector rojo es entonces un vector propio de la transformación, mientras que el azul no lo es. Dado que el vector rojo no ha cambiado de longitud, su valor propio es 1. Todos los vectores de esta misma dirección son vectores propios, con el mismo valor propio. Forman un subespacio del espacio propio de este valor propio.



Unos 20 años más tarde, Lagrange (José Luis de Lagrange; Turín, 25 de enero de 1736-París, 10 de abril de 1813) dio una versión más explícita de esta misma idea, al hallar la solución de un sistema de ecuaciones diferenciales, encontrando las raíces de lo equivalente a la ecuación característica de la matriz de coeficientes. El sistema particular de ecuaciones diferenciales surgió del examen de los "movimientos infinitesimales" de un sistema mecánico en la vecindad de su posición de equilibrio. En 1774, Lagrange resolvió un problema similar de mecánica celeste usando la misma técnica.

El nombre de **Cadenas de Markov** se debe al matemático ruso Andrei Andreevich Markov (1856-1922), quien las definió por primera vez en un

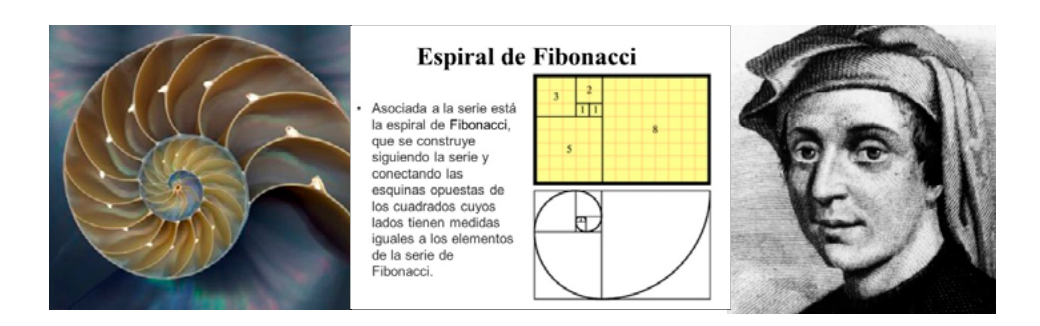


artículo de 1906 que trataba de la ley de los grandes números y posteriormente demostró muchos resultados estándar sobre ellas. Su interés en estas sucesiones se originó en las necesidades de la teoría de probabilidad. Markov nunca trató sus aplicaciones a las ciencias. Los únicos ejemplos reales que utilizó eran de textos literarios, donde los dos estados posibles eran vocales y consonantes. Para ilustrar sus resultados hizo un estudio estadístico de la alternancia de vocales y consonantes en el libro de Pushkin: Eugen Onegin.

Andrei Markov dio clases en la Universidad de San Petersburgo de 1880 a 1905, y se retiró para dar paso a matemáticos más jóvenes. Además de su trabajo en probabilidad, hizo contribuciones a campos como teoría de números, fracciones continuas y teoría de la aproximación. Fue un participante activo en el movimiento liberal ruso en la época anterior a la Primer Guerra Mundial; en varias ocasiones criticó públicamente la actuación de las autoridades estatales. En 1913, cuando como miembro de la Academia de Ciencias se le pidió participar en las pomposas ceremonias de celebración del 300 aniversario de la dinastía Romanov, prefirió organizar una conmemoración del 200 aniversario de la publicación de la ley de los grandes números de Jacobo Bernoulli.

FIBONACCI (1175-1230) El notable matemático italiano Leonardo de Pisa mejor conocido por su sobrenombre de Fibonacci, una abreviación de "filus Bonacci" que significa el hijo de Bonaccio de Pisa. Su padre, notario público de Bugía, en Argel, confió su hijo a las enseñanzas de un maestro de cálculo

musulmán: dedicó "algunos días" al estudio del ábaco, aprendió los signos numéricos de los indios, y penetró en seguida en los secretos de la ciencia arábiga. Ulteriores viajes lo llevaron posteriormente a través de Oriente y de Europa, hacia Egipto, Siria; Grecia, Sicilia y Provenza. Por doquier planteó nuevas cuestiones a los más afamados maestros con los que medía sus fuerzas para perfeccionar su sabiduría. "Pero todo esto, y el algoritmo y los arcos de Pitágoras, solo me parecieron otras tantas equivocaciones, comparadas con el método de los indios".



En el año 1202, a los 27 años de edad, fue cuando escribió Leonardo, resumiendo su sabiduría singular, un gran libro sobre el arte de calcular, el Liber abaci, en el que por primera vez un matemático cristiano ofrecía a sus contemporáneos una imagen completa, acabada y ordenada con la mayor claridad, del arte del cálculo tal y como éste se había desarrollado más allá del mundo europeo. Nos informa en ella que, habiendo aprendido en Berbería, el sistema numérico de los árabes lo transmitía "a la raza latina para que ésta no permaneciese más tiempo privada de él".

En el Liber abaci comenzaba Leonardo explicando los signos numéricos, enseñando a continuación a calcular con los dedos, así como las operaciones más sencillas con los números enteros y las fracciones. Una gran parte de sus problemas se referían a casos prácticos. Así, por ejemplo, en la sección 8 trataba "de los precios de las mercancías a largo plazo", en la 9, "del intercambio de mercancías y cosas semejantes"; en la 10, "de la aligación y compañía"; en la 11, "de la aleación y mezcla de las monedas"; en la 15, "de las reglas pertenecientes a la Geometría y de los problemas del Algebra y Almucabala". Es en éste mismo libro donde expone su ejemplo de los conejos y como se construye la sucesión que posteriormente llevará su nombre.



En 1826, **Cauchy** analizo las formas cuadráticas con tres variables, es decir, formas del tipo $Ax^2 + By_2 + Cz^2 + 2Dxy + 2Exz + 2Fyz$. Demostró que la ecuación característica formada a partir del determinante:

$$\begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix}$$

permanece igual bajo cualquier cambio de ejes rectangulares, lo que nosotros llamaríamos un cambio coordenado ortogonal.

7.11. EJERCICIOS PROPUESTOS

1. Propiedades de una matriz ortogonal Q.
   Sean Q una matriz ortogonal de $n \times n$ y x, y dos vectores columna ortogonales arbitrarios de R^n entonces:

   i) $(Qx)(Qy) = xy$; preservación del producto interior.
   ii) $\| Qx \| = \| x \|$; preservación de la longitud.
   iii) Ángulo entre xy = ángulo entre $QxQy$; preservación del ángulo.

   Demuestre las propiedades i) y ii).

2. Si Q_1 y Q_2 son matrices ortogonales, mostrar que también $Q_1\ Q_2$ es ortogonal. Si Q_1 es una rotación en Θ y Q_2 es una rotación en Φ ¿Qué es $Q_1\ Q_2$?

3. La factorización QR. Cualquier matriz A con columnas linealmente independientes puede factorizarse en un producto $A=QR$. Las columnas de Q son ortonormales y R es triangular superior e invertible. Si la matriz original A es cuadrada, también lo son sus factores Q y R, y entonces Q será una matriz ortogonal.

   a) Aplique el proceso de Gram-Schmidt a: $a_1=[0,0,1]^t$, $a_2=[0,1,1]^t$, $a_3=[1,1,1]^t$, y escriba el resultado en la forma $A=QR$.
   b) Factorice la matriz A en la forma QR.

$$A=\begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$$

4. Usando Easyfit o algún software de ajuste de datos, encuentre el ajuste lineal de los puntos siguientes:

   a) (1,3), (–2,4), (7,0).
   b) (1,2), (0,0), (2,7)
   c) (1,–3), (4,6), (–2,5), (3,–1).

5. Encuentre el mejor ajuste cuadrático a los puntos de datos:

   a) (2,–5), (3,0), (1,1), (4,–2).
   b) (–7,3), (2,8), (1,5).
   c) (1,–1), (3,–6), (5,2), (–3,1), (7,4).

6. Un fabricante compra grandes cantidades de ciertas refacciones para máquinas. Encuentra que su costo depende del número de cajas de piezas que compra una vez y que el costo por unidad disminuye al aumentar el número de unidades compradas. Supone que el costo es una función cuadrática del volumen y de experiencia anterior, obtiene la tabla siguiente:

No de cajas	Costo total
10	150
30	260
50	325
100	500
175	670

   Encuentre la función de costo total.

7. Usando el ejemplo que se adjunta conteste la siguiente pregunta ¿puede encontrarse una raíz cuadrada R de la matriz A de rotación de 90° y verificar que $R^2=A$? Recordar que R y R^2 comparten los mismos vectores propios, en otras palabras, la misma S. En realidad, hay 8 raíces cuadradas R posibles porque $\sqrt{1} = \pm 1$, $\sqrt{i} = \pm \frac{(1+i)}{\sqrt{2}}$, $\sqrt{-i} = \pm \frac{(1-i)}{\sqrt{2}}$.

8. a) Determine los valores y vectores propios de la matriz

$$A=\begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

Verifique que la traza de A es igual a la suma de los valores propios y que el determinante es igual al producto de valores propios.

b) Resuelva el sistema $u(t)=Au(t)$ si $u=(0, 0)^T$ ¿Cuáles son las dos soluciones exponenciales puras?

9. a) Si A y B comparten la misma matriz de vectores propios S de modo tal que $A=SAE_1S^{-1}$ y $B=SAE_2S^{-1}$ pruebe que $AB=BA$ (Nota: recuerde que $AE_1AE_2=AE_1E_2$ ya que se trata de matrices diagonales)

b) Si A está definida como sigue, encuentre un ejemplo para dicha B.

$$A=\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

10. Si cada número es el promedio de los dos números anteriores, $G_{k+2} = ½\ (G_{k+1} + G_k)$, construya la matriz A y diagonalícela comenzando con $G_0=0$ y $G_1=½$', encuentre una fórmula para G_k y calcule su límite cuando $k \to \infty$.

Bibliografía

1. Anton H., *Introducción al Algebra Lineal,* Ed. Limusa, 1984.
2. Fuentes Maya, S., *Apuntes de Matemáticas Aplicadas* I. DEPFI UNAM, 1989.
3. Fraleigh, J.B. y Beauregard, R.A. *Algebra Lineal.* Ed. Addison-Wesley Iberoamericana, 1989.
4. Grossman, S. I. *Aplicaciones de Algebra Lineal.* Ed. Grupo Editorial Iberoamérica, 1988.
5. Grossman, S. I. *Algebra Lineal.* Ed. Mc Graw Hill, 2008.
6. López de Medrano, S., *Modelos Matemáticos.* Ed. ANUIES, 1973.
7. López de Medrano, S., *Lenguajes Simbólicos.* Ed. ANUIES, 1973.
8. Perry, W. L. Algebra *Lineal con Aplicaciones.* Ed. Mc. Graw Hill, 1990.
9. Proskuriakov, I. *Problemas de Algebra Lineal.* Ed. MIR, 1986.
10. Rorres, C. y Anton, H. *Aplicaciones de Algebra Lineal.* Ed. Limusa, 1979.
11. Strang, G., *Algebra Lineal y sus Aplicaciones.* Ed. Fondo Educativo Interamericano, 1982.
12. Strang, G., *Introduction to Applied Mathematics.* Ed. Wellesley-Cambridge Press, 1986.

Enlaces

Capítulo 2
http://www.revistaciencias.unam.mx/es/86-revistas/revista-ciencias-67/747-un-vistazo-a-la-teoria-de-graficas.html.
http://www.revistaciencias.unam.mx/es/86-revistas/revista-ciencias-67/747-un-vistazo-a-la-teoria-de-graficas.html.
http://www.revistaciencias.unam.mx/es/86-revistas/revista-ciencias-67/747-un-vistazo-a-la-teoria-de-graficas.html

Capítulo 3
"Finite Markov Chains" por J. Kemeny y J. Snell, editado por Springer-Verlag, New York, Inc., 1976.

Capítulo 4
http://tochtli.fisica.uson.mx/electro/vectores/definici%C3%B3n_de_vectores.htm

Capítulo 5
Los espacios vectoriales, el amarillo, el rojo y el azul, Tomado de Suma 37, junio 2011, pp.75–82, Miguel Ángel Moreno Redondo



Álgebra lineal y sus aplicaciones, se publicó el 30 de abril de 2021
en la plataforma oficial de la Unidad de Apoyo Editorial (UDAE)
de la Facultad de Ingeniería, Ciudad Universitaria, México,
Ciudad de México. C.P. 04510

La familia tipográfica utilizada es Chivo para encabezados
y Minion Pro para texto, con sus respectivas variantes.